PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: METHODS OF USING NUCLEOTIDE INTEGRASE FOR CLEAVING DNA AND ATTACHING NUCLEIC ACIDS

(57) Abstract

The present invention provides new methods, employing a nucleotide integrase, for cleaving double-stranded and single stranded DNA substrates at specific sites and for attaching nucleic acid molecules to the cleaved DNA substrates. One method uses a nucleotide integrase to cleave one strand of a double-stranded DNA and to concomitantly attach a nucleic acid molecule to the cleaved strand. Another method uses a nucleotide integrase to cleave both strands of a double-stranded DNA substrate and to attach a nucleic acid molecule to one strand of the DNA substrate. Another method uses a nucleotide integrase to cleave both strands of a double-stranded DNA substrate and to attach an RNA molecule to one strand of the substrate and for attaching a cDNA to the other strand of the substrate. Another method cleaves single stranded DNA with the concomitant insertion of a nucleic acid molecule at the cleavage point. The nucleotide integrase comprises an RNP particle which comprises a group II intron RNA bound to a group II intron encoded protein. The present invention also relates to purified and reconstituted RNP particles and reconstituted RNP that cleave DNA substrates.

METHODS OF USING NUCLEOTIDE INTEGRASE FOR CLEAVING DNA AND ATTACHING NUCLEIC ACIDS

BACKGROUND

In recent years, a number of methods and biomolecules have been developed for manipulating DNA. Some of these biomolecules are used to cut or cleave DNA, which in some instances renders the substrate DNA nonfunctional. Other biomolecules are used to facilitate insertion of

5   new pieces of nucleic acid into the cleavage site of the DNA substrate. The insertion of new segments of nucleic acid into the cleavage sites of the DNA substrate changes the characteristics of the RNA or protein molecules encoded by the substrate DNA molecules. Accordingly, the biomolecules which catalyze the cleavage of DNA substrates or the

10   insertion of new nucleic acid molecules into the DNA substrates are useful tools for genetic engineering, for analytical studies and for diagnostic studies. One such molecule used for cleaving DNA substrates is the restriction endonuclease.

Restriction endonucleases are enzymatic proteins that cleave

15   double-stranded DNA. Such endonucleases recognize specific nucleotide sequences in double-stranded DNA, and cleave both strands within or near the specific recognition site. Such specificity renders the restriction endonucleases important tools in the controlled fragmentation of double-stranded DNA. Restriction endonucleases are

20   also useful analytical tools for determining whether certain sequences are present in substrate DNA and in genomic sequencing studies.

However, restriction endonucleases only cleave DNA substrates; they do not insert new nucleic acid molecules into the cleaved DNA substrate. Accordingly, another biomolecule is needed to insert new

25   pieces of DNA or RNA into the double-stranded DNA.

Ribozymes are catalytic RNA molecules that cleave RNA and, in certain circumstances, that insert new pieces of RNA into the cleavage site of the RNA substrate. Unfortunately, ribozymes have not been particularly useful for cleaving DNA. Ribozymes cut only single-

30   stranded DNA and only under extreme conditions of elevated temperatures and high concentrations of magnesium. The ribozymes have not been found to cleave double-stranded DNA. Ribozymes can be used to cleave double-stranded DNA only after the DNA is denatured and separated into two pieces of single stranded DNA.

35   Accordingly, it would be desirable to have a tool that cleaves double stranded DNA at specific sites and simultaneously inserts a new

-2-

nucleic acid molecule into the cleavage site of the double-stranded
DNA.

SUMMARY OF THE INVENTION

The present invention provides new methods, employing a
5   nucleotide integrase, for cleaving double stranded and single stranded
DNA substrates at specific sites and for inserting nucleic acid
molecules into the cleaved DNA substrate.   Thus, the nucleotide
integrase is a useful tool, particularly for genome mapping and for
genetic engineering.

10   One method uses a nucleotide integrase to cleave one strand of
double-stranded DNA at a specific site and to concomitantly attach a
nucleic acid molecule, which comprises an RNA molecule, to the cleaved
strand at the cleavage site.   Another method uses a nucleotide
integrase for cleaving both strands of double-stranded DNA and for

15 attaching a nucleic acid molecule, which comprises an RNA molecule, to
one strand of the DNA substrate. Preferably, the nucleic acid molecule
is fully integrated into the cleavage site. Another method uses a
nucleotide integrase for cleaving both strands of double-stranded DNA
and for attaching an RNA molecule to one strand of the DNA substrate at

20 the cleavage site and for attaching a cDNA to the other strand of the
DNA substrate at the cleavage site. Preferably, the RNA molecule is
fully integrated into the cleavage site of the one strand. Another
method provides for cleavage of single stranded DNA with the
concomitant insertion of a nucleic acid at the cleavage point.

25   The nucleotide integrase comprises a group II intron RNA and a
group II intron encoded protein, which is bound to the group II intron
RNA.   The nucleotide integrase includes at least two forms: a
ribonucleoprotein particle, also referred to herein as "RNP particle";
and a reconstituted particle.

30   The RNP particle comprises an excised group II intron RNA and a
group II intron-encoded protein, bound to the excised group II intron
RNA.   The excised group II intron RNA of the RNP particle has a
hybridizing region which is complementary to a recognition site on the
substrate DNA. Preferably,, the group II intron RNA has six domains.

35 The group II intron-encoded protein of the RNP particle preferably has
an X domain and a Zn domain. The X domain of the protein has a maturase
activity. The Zn domain of the protein has Zn^{2+} finger-like motifs.

The reconstituted particle is a synthetic particle that is
purified from a reconstituted RNP particle preparation.   The
40 reconstituted RNP particle preparation is made by combining a group II
intron RNA-protein complex, also referred to as the "RNA-protein
complex" and an exogenous nucleic acid, preferably an excised group II
intron RNA, hereinafter also referred to as "exogenous RNA". The RNA-
protein complex contains a group II intron-encoded protein that is

-3-

associated with RNA that is free of the excised group II intron RNA having a sequence which encodes the group II encoded protein. The intron-encoded protein has an X domain and a Zn domain. The exogenous RNA has a hybridizing region which is complementary to a recognition
5   site on the substrate DNA. Preferably, the exogenous RNA has six domains.

The present invention also relates to certain embodiments of the nucleotide integrase.

BRIEF DESCRIPTION OF THE DRAWINGS

10   Figure 1 is a diagram of the domains of the proteins which are encoded by the open reading frames of the group II intron 2 of the *S. cerevisiae* mitochondrial *COX1* gene, the group II intron 2 of the *M. polymorpha* mitochondrial *COX1* gene, and the group II intron 1 of the *N. tabacum* chloroplast *trnK* gene;

15 .   Figure 2 is a chart which depicts the locations of the mutations in the group II intron encoded proteins that are present in the
. nucleotide integrases of examples 4-9 and the RNP particles of comparative examples B-D;

Figure 3 is a diagram of the plasmid map of plasmid pJVM4.

20   DETAILED DESCRIPTION OF THE INVENTION

The present invention provides new methods that employ a nucleotide integrase for manipulating a DNA substrate. The methods enable double-stranded DNA substrates to be cleaved at specific sites and nucleic acid molecules to be inserted into the cleaved DNA
25 substrate. One method uses a nucleotide integrase to cleave one strand of double-stranded DNA at a specific site and to concomitantly attach a nucleic acid molecule, which comprises an RNA molecule, to the cleaved strand at the cleavage point. Another method uses a nucleotide integrase for cleaving both strands of double-stranded DNA and for
30 attaching a nucleic acid molecule, preferably an RNA molecule, to one strand of the DNA substrate at the cleavage point. Another method uses a nucleotide integrase to cleave both strands of double-stranded DNA and to attach a nucleic acid molecule, preferably an RNA molecule, to one strand of the DNA substrate and then to attach a cDNA molecule to
35 the other strand of the DNA substrate at the cleavage site.

The nucleotide integrase

The nucleotide integrase comprises a group II intron encoded RNA and a group II intron encoded protein which protein is bound to the RNA. The group II introns comprise a specific type of intron which is
40 present in the DNA of bacteria and in the DNA of organelles, particularly the mitochondria of fungi, yeast and plants and the chloroplast of plants. The group II intron RNA molecules, that is, the RNA molecules which are encoded by the group II introns, share a

-4-

similar secondary and tertiary structure.   The group II intron RNA
molecules typically have six domains. Domain IV of the group II intron
RNA contains the nucleotide sequence which encodes the "group II intron
encoded protein."   "Excised group II intron RNA," as used herein,
5   refers to the RNA that is either an _in vitro_ or _in vivo_ transcript of
the DNA of the group II intron and that lacks flanking exon sequences.
"Group II intron encoded protein" as used herein, is a protein encoded
by a group II intron.   The domains of three representative group II
intron-encoded proteins are depicted in Figure 1.
10   The nucleotide integrase includes RNP particles isolated from
wild type or mutant organisms and reconstituted RNP particles
synthesized from exogenous RNA and a particle preparation which lacks
the exogenous RNA.

The RNP particle

15   The RNP particle is used in an RNP particle preparation which,
while isolated from organelles, still contains substantial amounts of
ribosomes; or the RNP particle is used as a purified RNP particle which
is substantially free of ribosomes.
        The RNP particle and RNP particle preparation cleave both
20   strands of the double-stranded DNA,   catalyze the attachment of the
excised, group II intron RNA molecule to one of the strands of the
substrate DNA and catalyze the formation of a cDNA molecule on the
other strand of the cleaved double-stranded DNA substrate.
        The RNP particle   comprises an excised group II intron RNA and
25   a group II intron-encoded protein, which is bound to the excised group
II intron RNA.   The excised group II intron RNA of the RNP particle has
at least one hybridizing region, which will hybridize a complementary
recognition site on the substrate DNA.   The hybridizing region has a
nucleotide sequence, referred to hereinafter as the EBS sequence, which
30   is complementary to the sequence, referred to hereinafter as the IBS
sequence, of the recognition site of the intended substrate DNA. The
group II intron-encoded protein of the RNP particle preferably has an
X domain and a Zn domain. More preferably, the group II intron-encoded
protein further comprises a reverse transcriptase domain.
35   The excised group II intron RNA is obtained from wild type
organisms, or mutated organisms, by in vitro transcription, or by in
vitro transcription and splicing from the transcript of a modified or
unmodified group II intron. Nucleotide integrases in which the group
II intron RNA has nucleotide base changes in the hybridizing region, as
40   compared to the wild type, typically have altered specificity for the
substrate DNA of the nucleotide integrase. The group II intron RNA
also includes modified group II intron RNA molecules that have
nucleotide base changes or additional nucleotides in the internal loop

regions of the group II intron RNA, preferably the internal loop region
of domain IV.

The group II intron-encoded protein preferably has an X domain
and a Zn domain. The group II intron-encoded protein includes
5   proteins isolated from wild type organisms or from mutant organisms.
In addition to the conventional group II intron encoded proteins, other
proteins suitable as components of the nucleotide integrase include
modified group II intron encoded proteins which have additional amino
acids at the N terminus, or C terminus, or alterations in the internal
10   regions of the protein. Preferably, the group II-intron encoded
protein has a reverse transcriptase domain.

The Reconstituted particle

The reconstituted particle is used in a reconstituted RNP
particle preparation which, while isolated from organelles, still
15   contains substantial amounts of ribosomes; or the reconstituted
particle is used as a purified reconstituted RNP particle which is
substantially free of ribosomes.

The reconstituted particle is useful for cleaving one strand of
a double-stranded DNA substrate and attaching an RNA molecule to the
20   cleaved strand at the cleavage point; cleaving both strands of the
double-stranded DNA substrate and catalyzing the insertion of an RNA
molecule into the cleavage site of one strand of the DNA substrate;
cleaving both strands of a double-stranded DNA substrate and attaching
a nucleic acid molecule to both strands at the cleavage points.

25   The reconstituted RNP particle preparation comprises an
exogenous excised group II intron RNA, an RNA-protein complex, and a
reconstituted RNP particle which comprises the exogenous group II
intron RNA bound to a group II intron-encoded protein.

The RNA-protein complex contains a group II intron-encoded
30   protein that is associated with RNA that is free of the excised, group
II RNA molecules having a sequence which encodes the group II intron-
encoded protein. The group II intron-encoded protein comprises an X
domain and a Zn domain and includes both unmodified and modified group
II intron-encoded proteins. The X domain of the protein has a maturase
35   activity. The Zn domain of the protein has Zn^{2+} finger-like motifs.
More preferably, the group II intron-encoded protein further comprises
a reverse transcriptase domain. The group II intron-encoded protein
has a binding affinity for the RNA of the complex and for the exogenous
excised group II intron RNA.

40   The exogenous group II intron RNA of the RNP particle has at
least one hybridizing region which is complementary to a recognition
site on the substrate DNA. The exogenous RNA preferably has six
domains. The exogenous RNA includes both unmodified and modified group
II intron RNA molecules. Preferably, the exogenous RNA is an _in vitro_

transcript or a derivative of an in vitro transcript of an unmodified
or modified intron group II intron. For example, the exogenous RNA may
be derived by splicing from an in vitro transcript. In a preferred
embodiment of the reconstituted particle, the exogenous RNA has the
5   sequence encoded by the DNA sequence of SEQ.ID.NO.1 and the group II
intron-encoded protein is a 62 kDa protein that is processed from a
polypeptide having the sequence of SEQ. ID. NO. 2.

Methods for Cleaving DNA and Inserting Nucleotides

        The methods of the present invention and the reconstituted RNP
10 particles are useful analytical tools for determining the location of
a defined sequence in a double-stranded DNA substrate. Moreover, the
simultaneous insertion of the nucleic acid molecule into the first
strand of DNA permits tagging of the cleavage site of the first strand
with a radiolabeled molecule. In addition, the automatic attachment of
15 an RNA molecule onto one strand of the DNA substrate permits
identification of the cleavage site through hybridization studies that
use a probe that is complementary to the attached RNA molecule. An
attached RNA molecule that is tagged with a molecule such as biotin
also enables the cleaved strand to be affinity purified. These methods
20 are also useful for rendering the substrate DNA nonfunctional or for
changing the characteristics of the RNA and protein encoded by the
substrate DNA.

        While a wide range of temperatures are suitable for the methods
herein, good results are obtained at a reaction temperature of from
25 about 30°C to about 42°C, preferably from about 30° to about 37°C. A
suitable reaction medium contains a monovalent cation such as Na⁺ or K⁺,
at a concentration from about 0 to about 300 mM; preferably from about
50 to about 200 mM KCl, and a divalent cation, preferably a magnesium
or manganese ion, more preferably a magnesium ion, at a concentration
30 that is less than 100 mM and greater than 1 mM. Preferably the
divalent cation is at a concentration of about 5 to about 20 mM, more
preferably about 10 to about 20 mM. The rate of attachment of the
intron RNA to the cleaved sense strand at 5 mM $MgCl_2$ is about 25% of the
rate that is obtained using 10-20 mM $MgCl_2$. The preferred pH for the
35 medium is from about 6.0-8.5, more preferably about 7.5-8.0.

Cleaving one strand of double stranded DNA

        Double stranded DNA substrate having a recognition site is
cleaved and RNA inserted at the cleavage site by reacting the substrate
DNA with a nucleotide integrase. Suitable nucleotide integrases
40 include : a mitochondrial RNP particle from mutant yeast strain $lo2^{HHVR}$
which comprises a modified, excised RNA from the group II intron aI2 of
the yeast mitochondrial COX1 gene and an aI2-encoded protein that has
a missense mutation in the HHVR motif; a mitochondrial RNP particle
from mutant yeast strain $l°2^{ConZn}$ which comprises a modified, excised aI2

intron RNA and an aI2-encoded protein that lacks the most conserved motifs in the Zn domain; a mitochondrial RNP particle from mutant yeast strain $1°2^{c-c/2}$ that comprises a modified, excised aI2 intron RNA and an aI2-encoded protein that has a missense mutation in the second Zn^{+2}

5   finger-like motif.

The most preferred reaction conditions for attachment of the intron RNA to the cleaved strand of substrate DNA are 100 mM KCl, 20 mM $MgCl_2$, pH 7.5, 5 mM DTT and 37°C.

<u>Cleaving both strands of double stranded DNA</u>

10   Both strands of a DNA substrate having a recognition site are cleaved by reacting nucleotide integrases with the DNA.   Suitable nucleotide integrases include, for example: a mitochondrial RNP particle that comprises an excised aI2 RNA and aI2-encoded protein, a mitochondrial RNP particle that comprises aI1 RNA and aI1 intron-

15 · encoded protein; a mitochondrial RNP particle from mutant yeast strain $1°2^{YAHH}$ which comprises a modified, excised aI2 RNA and an aI2-encoded
· protein that has the mutation YADD→YAHH in the reverse transcriptase domain of the protein; and a reconstituted RNP particle which comprises an exogenous excised aI2 RNA and an aI2-encoded protein.   The most

20   preferred reaction conditions for cleavage of both strands of the DNA substrate by the nucleotide integrase, are 100 mM KCl, 20 mM $MgCl_2$, pH 7.5, 5 mM DTT and 37°C.   Under these conditions, the nucleotide integrase which comprises an excised aI1 RNA and an aI1 encoded protein cleaves and inserts the aI1 RNA into the cleavage site of the sense

25   strand of a double-stranded DNA substrate which contains the target sequence of SEQ. ID. NO. 17, in the sense strand.   In this sequence, nucleotide bases 63-69 having the sequence 5'TTAATG, hereinafter designated as "IBS1," are complementary to nucleotide bases, hereinafter referred to as "EBS1," in domain I of the aI1 intron RNA.

30   Nucleotide bases 57-62 which have the sequence 5'CAGTTA, hereinafter designated as "IBS 2," are complementary to other nucleotide bases, hereinafter referred to as "EBS2," in domain I of the aII intron RNA. The RNP particles of this nucleotide integrase cleave the sense strand of this substrate between nucleotide 69 and nucleotide 70 of the target sequence, and cleave the antisense strand of this substrate 10 base

35   sequence, and cleave the antisense strand of this substrate 10 base pairs downstream from cleavage site of the sense strand.

Under these conditions, a nucleotide integrase which comprises an excised aI2 RNA and an aI2 encoded protein cleaves and inserts the aI2 RNA into the cleavage site of the sense strand of a double-stranded

40   DNA   substrate   which   contains   the   target   sequence 5'TTTTAGTAGCTGGTCATGCTGTATTAATAATTTTCTTCTTAGTAATGCCTGCTTTAATAG GAGGTTTTGGT), SEQ. ID. NO. 5, in the sense strand.   In this sequence, nucleotide bases 31-36 having the sequence 5'TTTTCT, hereinafter designated as "IBS3," are complementary to nucleotide bases,

hereinafter referred to as "EBS3," in domain I of the aI2 intron RNA. Nucleotide bases 24-30, which have the sequence 5'TTAATAA, hereinafter designated as "IBS4," are complementary to other nucleotide bases, hereinafter referred to as "EBS4," in domain I of the aI2 intron RNA.

5   The RNP particles of this nucleotide integrase cleave the sense strand of this substrate between nucleotide 36 and nucleotide 37 of the target sequence, and cleave the antisense strand of this substrate 10 base pairs downstream from cleavage site of the sense strand.

Cleaving DNA, insertion of RNA and formation of cDNA

10   Both strands of a DNA substrate having a recognition site are cleaved and an RNA attached to one strand of the DNA and a cDNA attached to the cleaved second strand of the DNA substrate by incubating the DNA with the nucleotide integrase in the presence of deoxynucleotides. Suitable nucleotide integrases include, for example:

15   a mitochondrial RNP particle that comprises an excised aI2 RNA and aI2- encoded protein and a mitochondrial RNP particle that comprises aI1 RNA and aI1 intron-encoded protein.

       Suitable conditions for the nucleotide integrase-catalyzed synthesis of a cDNA on the second strand of the DNA substrate include

20   a reaction temperature of from about 30° to about 45°C, preferably from about 37° to about 42°C. The preferred reaction medium includes a monovalent cation such as Na$^+$ or K$^+$, at a range of 0-300 mM, preferably from about 50 mM to about 150 mM KCl, and a divalent cation, preferably a magnesium or manganese ion, more preferably a magnesium ion at a

25   concentration range of from about 2 mM to about 20 mM. The preferred pH for the medium is about 7.5 to about 8.5. The most preferred conditions for synthesis of an elongated cDNA product are 100 mM KCl, 2 mM MgCl$_2$, 5 mM DTT, pH 8.5, and 37°C. The most preferred conditions for maximum cleavage of the second strand and the formation of a cDNA

30   product is 100 mM KCl, a Mg^{2+} concentration of 10 to 20 mM, 5 mM DTT and a pH of 7.5.

Cleavage of single stranded DNA

       Single stranded DNA substrate having a recognition site is cleaved and RNA inserted at the cleavage site by reacting the substrate

35   DNA with a nucleotide integrase. Suitable nucleotide integrases include for example : a mitochondrial RNP particle that comprises an excised aI2 RNA and aI2-encoded protein and a mitochondrial RNP particle that comprises aI1 RNA and aI1 intron-encoded protein.

       The most preferred reaction conditions for cleavage of the

40   substrate and insertion of the intron RNA into the cleavage site by the nucleotide integrase, are 100 mM KCl, 20 mM MgCl$_2$, pH 7.5, 5 mM DTT and 37°C. Under these condition, the nucleotide integrase which comprises an aI2 intron RNA and an aI2 encoded protein cleaves and inserts an aI2 RNA into the cleavage site of a single-stranded DNA substrate that

                                -9-

contains the IBS3 sequence 5'TTTTCT or the IBS4 sequence 5'TTAATAA. In
contrast, the aI2 RNA alone is unable to cleave double stranded DNA or
single stranded DNA at the 20 mM Mg^{2+} concentration typically used in
the methods of the present invention.

5   Preparation of the Nucleotide Integrase
        The nucleotide integrase is isolated from wild type or mutant
yeast   mitochondria,   fungal   mitochondria,   plant   mitochondria,
chloroplasts,   the   proteotobacterium   Azotobacter   vinelandii,   the
cyanobacterium Calothrix, and Escherichia coli. The procedure for
10   isolating the RNP particle preparation involves mechanically and/or
enzymatically disrupting the cell membranes and or cell walls of the
organisms.   In the case of fungi and plants, the purification also
involves separating the specific organelles, such as mitochondria or
chloroplasts, from the other cellular components by differential
15   centrifugation   and/or   flotation   gradients   and   then   lysing   the
organelles with a nonionic detergent, such as Nonidet P-40.   The
organelle and bacterium lysates are then centrifuged through a sucrose
cushion to obtain the ribonucleoprotein (RNP) particle preparation.
The RNP particles may be further purified by separation on a sucrose
20   gradient,   or   a   gel   filtration   column,   or   by   other   types   of
chromatography.
        The reconstituted RNP particle preparation is prepared by
combining an RNA-protein complex preparation with an exogenous excised
group II intron RNA.   The RNA-protein complex is preferably isolated
25   from a yeast, fungi, or bacterium using the protocol for RNP particles
described above. The RNA-protein complex comprises a group II intron-
encoded protein and RNA that is free of the excised group II intron RNA
having a sequence that encodes the group II intron-encoded protein.
        The exogenous RNA preferably is made by _in vitro_ transcription
30   or by _in vitro_ transcription and self-splicing of the group II intron.
The exogenous RNA may also be made by isolation of the group II intron
RNA from cells or organelles in which it is naturally present or from
cells in which an altered intron has been inserted and expressed.   The
exogenous RNA is then added to a preparation containing the RNA-protein
35   complex.   Preferably, the exogenous group II intron RNA is first
denatured.   The exogenous RNA is added to the RNA-protein complex on
ice.

Examples
        The RNP particle preparations of the following examples 1-10,
40   the RNA-protein complex of the example 12, and the preparations of
comparative examples A-D were isolated from the mitochondria of the
wild-type Saccharomyces cerevisiae yeast strain ID41-6/161 MATa adel
lys1, hereinafter designated "161", and derivatives thereof.   The

-10-

mitochondria of the wild-type yeast strain 161 contains a *COX1* gene that includes the group II intron aI1 and the group II intron aI2.

The *COX1* gene in the mutant yeast strains either lacks one of the group II introns or has a mutation in one of the group II introns.
5   The excised group II intron RNA molecules and the group II intron encoded proteins are derived from the group II introns aI1 and aI2 that are present in the wild-type and mutant yeast strains.

Figure 2 depicts several of the mutations/modifications that are present in the aI2-encoded proteins that are produced by the different
10 mutant yeast strains. The intron composition of the *COX1* gene in the different yeast strains is denoted by a convention in which a superscript "+" indicates the presence of the aI1 intron or the aI2 intron, a superscript "o" indicates the absence of the aI1 or aI2 intron, and other superscripts refer to specific alleles or mutations
15 in the aI2 intron.

Example 1

An RNP particle preparation was isolated from the mitochondria of the *Saccharomyces cerevisiae* wild-type yeast strain 161. The intron composition of the *COX1* gene of the wild-type strain is 1^+2^+. The RNP
20 particle preparation contains an RNP particle that is derived from the aI1 intron and includes an excised, aI1 RNA bound to a protein encoded by aI1. The preparation also contains an RNP particle that is derived from the aI2 intron and that comprises a excised aI2 RNA molecule and an associated aI2-encoded protein.
25   To prepare the RNP particle preparation, the yeast were inoculated into a 1 liter liquid culture medium containing 2% raffinose, 2% BactoPeptone from Difco and 1% yeast extract from Difco to an $O.D._{595}$ of 1.6-1.7. The cell walls were digested with 40 mg of the yeast lytic enzyme from ICN, and the cells broken by mechanical
30 disruption with glass beads. The nuclei and cell debris were pelleted from the lysate by centrifugation for 5 minutes in a Beckman GSA rotor at 5,000 rpm. The supernatant was removed and centrifuged in a Beckman GSA rotor at 13,000 rpm for 15 minutes to obtain a mitochondrial pellet. The mitochondria were layered on a flotation gradient
35 consisting of a 44% sucrose solution layer, a 53% sucrose solution layer, and a 65% sucrose solution layer and centrifuged in a Beckman SW28 rotor at 27,000 rpm for 2 hours and 10 minutes. The mitochondria were collected from the 53%/44% interface and suspended in buffer containing 0.5M KCl, 50 mM $CaCl_2$, 25 mM Tris-HCl, pH 7.5, 5mM DTT and
40 lysed by the addition of Nonidet P-40 to a final concentration of 1%. The mitochondrial lysate was then centrifuged in a Beckman 50Ti rotor at 50,000 rpm for 17 hours through a 1.85 M sucrose cushion in a buffer containing 0.5M KCl, 25 mM $CaCl_2$, 25 mM Tris-HCl, pH 7.5, 5mM DTT, to obtain a pellet of RNP particles that were largely free of

mitochondrial proteins. The isolated RNP particles were resuspended in
10 mM Tris-HCl, pH 8.0 and 1 mM DTT and stored at
-70°C. The preparation may be repeatedly thawed and frozen before use.
Example 1a Purified RNP particle

5   2.5 O.D.$_{260}$ of the RNP particles from example 1 in a volume of
150 μl were layered onto a 12 ml 5-20% linear sucrose gradient in a
buffer consisting of 100 mM KCl, 2 mM MgCl$_2$, 50 mM Tris-HCl, pH 7.5, and
5 mM DTT. The gradient was centrifuged in an SW41 rotor at 4°C at
40,000 rpm for five hours. The gradient was fractionated into 35
10 fractions of approximately 325 μl. Fractions 12-20 contain the
purified RNP particles which are substantially free of ribosomal RNA.
The location of the RNP particles in the gradient fractions was
independently verified by Northern hybridization with aI2 antisense
RNA. The location of the small and large subunits of ribosomal RNA in
15 · the gradient fractions was independently verified by ethidium bromide
staining of the fractions on a 1% agarose gel.
· Example 2 RNP particle preparation from mutant yeast strain 1°2˙

       The RNP particles comprise an excised aI2 RNA and an aI2-encoded
protein. Yeast strain 1°2^{+t} was obtained from Dr. Philip S. Perlman at
20 the University of Texas Southwestern Medical Center and was prepared as
described in Moran et al., 1995, Mobile Group II Introns of Yeast
Mitochondrial DNA Are Novel Site-Specific Retroelements, Mol. Cell
Biol. 15, 2838-38, which is incorporated herein by reference. The 1°2^{+t}
mutant strain was constructed as follows: (i) the aI2 intron from
25 strain 161 was cloned as a ClaI-to-BamHI fragment into pBluescript KS˙
obtained from Stratagene to yield pJVM4; (ii) pJVM4 was cleaved with
ClaI and NdeI to remove the 5' end of the insert; and (iii) an MspI-to-
NdeI fragment that contains exons 1 and 2 of the mitochondria COX1 gene
plus the 5' end of aI2 from yeast strain C1036ΔI was inserted to yield
30 plasmid pJVM164. Yeast strain C1036ΔI, in which aI1 is excised from
the mitochondrial DNA, was prepared as described in Kennell et al.,
1993, Reverse transcriptase activity associated with maturase-encoding
group II introns in yeast mitochondria. Cell 73, 133-146, which is
incorporated herein by reference. pJVM164 was transformed into a [rho⁰]
35 strain, and the 1°2^{+t} allele was placed into an intact mitochondrial DNA
by recombination. This last step is accomplished by mating to a
nonreverting COXI mutant derived from mutant C1036 (strain 5B), whose
construction is described in Kennel et al., 1993, and selecting for
recombinant progeny that are capable of respiring and growing on
40 glycerol-containing medium (GLY˙) and that contain the transformed COXI
allele in place of the 5B allele.
       The reactions and manipulations directed at cloning DNA, such as
ligations, restriction enzyme digestions, bacterial transformation, DNA
sequencing etc. were carried out according to standard techniques, such

-12-

as those described by Sambrook et al., Molecular cloning: a laboratory
manual, 2nd ed., Cold Spring Harbor Laboratory Press. Cold Spring
Harbor. N.Y. Yeast mitochondrial transformations were also carried out
according to standard techniques such as those described in Belcher et
5   al., 1994, Biolistic transformation of mitochondria in Saccharomyces
cerevisiae, 101-115. In N.-S. Yang and P. Christou (ed.) Particle
Bombardment Technology for Gene Transfer. Oxford University Press, New
York. The RNP particle preparation was made from the mitochondria of
mutant yeast strain 1°2', as in Example 1.

10 Example 3 RNP particle preparation from mutant yeast strain 1⁺2°

Yeast strain 1⁺2° is a derivative of the wild-type yeast strain
161. The yeast strain 1⁺2° was obtained from Dr. Philip S. Perlman and
was prepared as described in Kennell et al., 1993. Cell 73, 133-146.
Yeast strain 1⁺2° contains a segment of the COX1 gene of S.
15 diastaticus, which lacks aI2, inserted into wild-type 161 mtDNA via
mitochondrial transformation. The construction started with plasmid
pSH2, which contains aI1 from wild-type 161 and some flanking sequences
cloned as a HpaII/EcoRI fragment in pBS' (Stratagene, La Jolla, CA).
That plasmid was cleaved near the 3' end of aI1 with ClaI and in the
20 downstream polylinker with BamHI, and the gap was filled with a
ClaI/BamHI fragment from S. diastaticus mitochondrial DNA (NRRL Y-2416)
that contains the 3' end of aI1, E2, E3 and most of aI3, thus creating
a 1°2° form of the COX1 gene. The plasmid containing the hybrid COX1-
1°2° segment was transformed into a ρ° derivative of strain MCC109 (MATα
25 ade2-101 ura3-52 karl-1) by biolistic transformation. The resulting
artificial petite was crossed to strain n161/m5B, and gly' recombinants
containing the COX1 1°2° allele in the n161 background were isolated.
The hybrid aI1 allele, which is spliced normally, differs from that of
wild-type 161 by one nucleotide change, C to T, at position 2401,
30 changing Thr₇₄₄ to Leu in the intron open reading frame. The RNP
particle preparation was made from the mitochondria of mutant yeast
strain 1°2° as in Example 1. The RNP particles comprise an excised aI1
RNA molecule and an aI1 encoded protein.

Example 4 RNP particle preparation from mutant yeast strain 1°2ᵞᴬᴴᴴ

35 Yeast strain 1°2ᵞᴬᴴᴴ was obtained from Dr. Philip S. Perlman and
was made as described in Moran et al., 1995, Mol. Cell Biol. 15, 2838-
38., using a mutagenized pJVM164 plasmid. The allele was made by
oligonucleotide-directed mutagenesis of pJVM164 which contains a 4.4 kb
MspI/BamHI fragment extending from 217 nucleotides upstream of exon 1
40 through intron aI3 of the COX1 allele. The mutagenesis changes the aI2
nucleotides 1473 to 1478 from GAT GAT to CAT CAT (D-491D-492 to HH).
The RNP particles comprise a mutated, excised aI2 RNA and an aI2-
encoded protein that has the mutation YADD→YAHH in the reverse
transcriptase domain of the protein. The RNP particle preparation was

made from the mitochondria of mutant yeast strain $1°2^{YAHH}$ as in Example
1.

Example 5 RNP particles from the mutant yeast strain $1°2^{P714T}$

     The mutant yeast strain $1°2^{P714T}$ was obtained from Dr. Philip S.
Perlman and was constructed according to the procedure described in
Kennell et al., 1993, _Cell_ 73, 133-146, where it is named n161/m161-
C1036Δ1. The RNP particles comprise a mutated, excised aI2 intron RNA
molecule and an aI2-encoded protein that carries the missense mutation
$P_{714}T$ in the Zn domain. The RNP particle preparation was made from
mitochondria of mutant yeast strain $1°2^{P714T}$ as in Example 1.

Example 6 RNP particle from mutant yeast strain $1°2^{HHVR}$

     The mutant yeast strain $1°2^{HHVR}$ was obtained from Dr. Philip S.
Perlman and was made as described in Moran et al., 1995, Mol. Cell
Biol. 15, 2828-38, which is incorporated herein by reference, using a
mutagenized pJVM164 plasmid. The allele was constructed by site-
directed mutagenesis of pJVM164. The aI2 intron has the following
changes: positions 2208-2219 from CATCACGTAAGA to GCAGCTGCAGCT
($H_{736}H_{737}V_{738}R_{739}$ to AAAA) and A_{2227} A to T ($N_{742}I$). This nucleotide integrase
preparation comprises a mutated, excised aI2 intron RNA and an aI2-
encoded protein that has a missense mutation in the HHVR motif. The
RNP particle preparation was made from mitochondria of mutant yeast
strain $1°2^{HHVR}$.

Example 7 RNP particle from mutant yeast strain $1°2^{ConZn}$

     The mutant yeast strain $1°2^{+ConZn}$ was obtained from Dr. Philip S.
Perlman and was made as described in Moran et al., 1995, Mol. Cell
Biol. 15, 2828-38, using a mutagenized pJVM164 plasmid. The allele was
constructed by oligonucleotide-directed mutagenesis of pJVM164. The
aI2 intron has the following changes:positions 2157-2165 changed from
TTATTTAGT to TAATAATAA ($L_{719}F_{720}S_{721}$ to OchOchOch). RNP particles comprise
a mutated, excised aI2 intron RNA and an aI2-encoded protein that lacks
the most conserved motifs in the Zn domain. The RNP particle
preparation was made from mitochondria of mutant yeast strain $1°2^{ConZn}$.

Example 8 RNP particle from mutant yeast strain $1°2^{C-C/1}$

     The mutant yeast strain $1°2^{C-C/1}$ was obtained from Dr. Phillip S.
Perlman and was made as described in Moran et al., 1995, using a
mutagenized pJVM164 plasmid. The allele was constructed by site-
directed mutagenesis of pJVM164. The aI2 intron has the following
changes:positions 2172-2173 changed from TG to GC ($C_{724}A$) and 2180-2182
changed from TTG to AGC ($I_{726}C_{727}$ to MA). The RNP particles comprise a
mutated, excised aI2 intron RNA and an aI2-encoded protein that has a
missense mutation in the first Zn^{+2}-finger-like motif. The RNP particle
preparation was made from mitochondria of mutant yeast strain $1°2^{C-C/1}$.

Example 9 RNP particles from mutant yeast strain $1°2^{C-C/2}$

-14-

The mutant yeast strain $1°2^{c-c/2}$ was obtained from Dr. Philip S. Perlman and was made as described in Moran et al., 1995, using a mutagenized pJVM164 plasmid. The allele was constructed by site-directed mutagenesis of pJVM164. The aI2 intron has the following
5   changes: position 2304-2305 changed from TG to GC $(C_{768}A)$ and 2313-2314 changed from TG to GC $(C_{771}A)$. The RNP particles comprise a mutated excised aI2 intron RNA and an aI2-encoded protein that has a missense mutation in the second Zn^{+2} finger-like motif. The RNP particle preparation was made from mitochondria of mutant yeast strain $1°2^{c-c/2}$.

10 Example 10 RNP particles from mutant yeast strain $1°2^{H6}$

The mutant yeast strain, obtained from Dr. Philip S. Perlman, was made by transferring the mutagenized plasmid pJVM164 into the mitochondria of yeast strain GRF18 as described in Moran et al., 1995. The allele was constructed by site directed mutagenesis of pJVM164 and
15 has the sequence CATCATCATCATCATCAT inserted between nucleotides 2357 and 2358 of the aI2 intron. The RNP particle preparation was made from mitochondria of mutant yeast strain $1°2^{H6}$ according to the protocol described above for example 1. The RNP particles comprise a mutated, excised aI2 intron RNA and an aI2-encoded protein that has six
20 histidines added to the C terminus of the aI2-encoded protein.

Example 11 RNP particles from Neurospora intermedia.

Mitochondria from the Varkud strain of Neurospora intermedia, which is available from the Fungal Genetics Stock Center, were prepared as described in Lambowitz A.M. 1979, Preparation and analysis of
25 mitochondrial ribosomes. Meth. Enzymol. 59, 421-433.. The conidia were disrupted with glass beads and the mitochondria and RNP particles isolated as described in Example 1. The RNP particles comprise an excised *coI* intron RNA and the protein encoded by the *coI* intron.

Example 12 Reconstituted RNP particle preparation

30 A reconstituted RNP particle preparation was made by incubating an exogenous, excised, *in vitro* RNA transcript of the aI2 intron with an RNP preparation isolated from the mutant yeast strain $1°2^{ΔD5}$, in which the aI2 intron RNA lacks a domain V and is therefore splicing defective. The mutant allele $1°2^{ΔD5}$ was obtained from Dr. Philip S.
35 Perlman and was constructed using the same procedure that was used to make yeast strain $1°2^{ΔD5}$ that was described in Moran et al. 1995, except that the final mating was with yeast strain $1°2˙$. The RNP preparation was isolated from $1°2^{ΔD5}$ using the protocol described above in Example 1. The RNP preparation isolated from the mitochondria of $1°2^{ΔD5}$ does not
40 contain excised aI2 RNA but does contain aI2-encoded protein that is associated with other RNA molecules in the preparation.

The exogenous RNA was made by *in vitro* transcription of the plasmid pJVM4 which includes a fragment of the yeast mitochondrial *COX1* gene from the *ClaI* site of the group II intron 1 (aI1) to the *BamHI*

site of aI3 that has been inserted into the pBLUESCRIPT KS+ plasmid. As shown in Figure 3, which depicts the plasmid map of pJVM4, plasmid pJVM4 contains the following COX1 sequences: Exon 2, aI2, Exon 3 and parts of aI1 and aI3 sequence. The sequences are operably linked to a

5 T3 RNA polymerase promoter. The Exon 2 and Exon 3 sequence are required for self-splicing of the aI2 intron RNA from the RNA transcript. pJVM4 was linearized with BstEII, which cuts at the 3' end of Exon 3 then 3 μg of the plasmid was incubated in 300 μl of 40 mM Tris-HCl at pH 8.0, 25 mM NaCl, 8 mM MgCl$_2$ 2 mM spermidine, 5 mM DTT 200

10 μM rNTPs, 182 U of RNasin from US Biochemical and 750 U of T3 RNA polymerase from BRL at 37°C for 2 hours to make the RNA transcripts. Following the incubation, the RNA transcripts were phenol extracted, purified on G-50 column, phenol extracted and precipitated with ethanol. The RNA transcripts were then incubated in 40 mM Tris-HCl at

15 pH 7.5, 100 mM MgCl$_2$, 2 M NH$_4$Cl at 40-45°C for 1 hour to allow self-splicing of the aI2 intron RNA molecules from the RNA transcripts and to obtain the splicing products. The splicing products, which include the excised aI2 RNA transcript, the ligated transcript which lacks the aI2 intron RNA, and the unspliced transcript, were desalted by passing

20 through a G-50 column, then phenol extracted and ethanol precipitated to provide the exogenous RNA. The exogenous RNA was then resuspended to a final concentration of 1.0 μg/μl in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA. A portion of the exogenous RNA was denatured by heating to 90°C for 2 minutes and then immediately put on ice.

25   To prepare the reconstituted RNP particle preparation, 1 μl of the denatured exogenous RNA was added to 2 μl of the 1°2^{ad5} RNP preparation (0.025 O.D.$_{260}$ units) on ice for 2 minutes. The preparation, was used immediately.

Example 12a Purified Reconstituted Particles

30   Reconstituted particles are purified from reconstituted preparation of Example 12 according to the method of 1a.

Example 13 RNP Particle Preparation

          An RNP particle preparation in which the group II intron RNA of the RNP particle has a "modified" EBS sequence that is complementary to

35 a selected sequence on an intended single strand DNA substrate is prepared by two methods. First, oligonucleotide-directed mutagenesis of the group II intron DNA is performed by standard, well-known methods to change the nucleotides which encode the EBS sequences of the group II intron RNA. These sequences are in domain I of the group II intron

40 RNA. The mutagenized intron DNA is then inserted into a vector, such as a plasmid, where it is operably linked to an RNA polymerase promoter, such as a promoter for T7 RNA polymerase or SP6 RNA polymerase or T3 RNA polymerase and an in vitro transcript of the modified group II intron RNA is made as described in example 12. The modified exogenous

-16-

RNA transcript is then combined with an RNA-protein complex that has been isolated as described in example 12 to provide a modified reconstituted RNP particle preparation.

5   Alternatively, an RNP particle preparation in which the EBS sequences of the group II intron RNA are modified is prepared by site-directed mutagenesis of an organism, such as a yeast, as described in examples 3-9, and by isolation of the modified RNP particle preparation from the organism as described in example 1.

Example 13a Purified Modified EBS Reconstituted Particles

10   RNP particles in which the EBS sequences of the group II intron RNA have been modified are purified from the preparation of Example 13 according to the method of 1a.

Example 14

An RNP particle preparation containing an RNP particle in which
15   the loop region of domain IV of the group II intron RNA is modified, ·
that is the loop region nucleotide sequence of domain IV differs from the nucleotide sequence of the aI2 RNA of Examples 1-10 is prepared by two methods. First oligonucleotide-directed mutagenesis of the aI2 intron DNA is performed by standard, well-known methods to change the
20   nucleotide sequences which encode for the loop region of domain IV of the aI2 intron RNA. The mutagenized aI2 intron DNA is then inserted into a vector, such as a plasmid, where it is operably linked to an RNA polymerase promoter, such as a promoter for T7 RNA polymerase or SP6 RNA polymerase or T3 RNA polymerase and an in vitro transcript of the
25   modified group II intron RNA made as described above in example 12. The exogenous RNA is then combined with an RNA-protein complex that has been isolated as described for example 12 to produce a modified reconstituted RNP particle preparation.

Alternatively, an RNP particle preparation in which the EBS
30   sequences of the group II intron RNA are modified is prepared by site-directed mutagenesis of an organism, such as a yeast, as described in examples 4-10, and by isolation of the RNP particle preparation from the organism as described in example 1. IV is performed as in examples 4-10 of the aI2 intron DNA.

35   Example 14a Purified Reconstituted Particles

Modified RNP particles are purified from the preparation of Example 14 according to the method of 1a.

Comparative Examples

Comparative Example A

A ribonucleoprotein preparation was isolated from the mitochondria of the mutant yeast strain $1°2°$, according to the protocol described in Example 1 to produce Comparative Example A. The mitochondrial COX1 gene of mutant yeast strain $1°2°$ lacks the aI1 and aI2 introns. Thus, the RNP preparation of Comparative Example A lacks excised aI1 RNA, excised aI2 RNA, and the proteins encoded by aI1 and aI2.

Comparative Example B

An RNP preparation was isolated from mutant yeast strain $1°2^{\Delta DS}$ according to the protocol described in Example 1 to produce Comparative Example B. The RNP preparation of Comparative Example B contains an aI2-encoded protein but lacks excised aI2 RNA.

Comparative Example C

An RNP preparation was isolated from the mitochondria of mutagenized yeast strain $1°2^{X4}$ according to the protocol described in Example 1 to produce Comparative Example C. The mutant allele $1°2^{X4}$ was obtained from Dr. Philip S. Perlman. The mutant allele was constructed by site-directed mutagenesis of pJVM16 and has the following changes in the aI2 intron: positions 2004-2012 changed from AAATTCGGT to GCAGCTGCT ($K_{668}F_{669}G_{670}$ to AAA). The RNP preparation of comparative example C contains a modified, excised aI2 RNA and an aI2-encoded protein that has a missense mutation in domain X.

Comparative Example D

An RNP preparation was isolated from the mitochondria of mutagenized yeast strain $1°2^{\Delta Zn}$ to produce comparative Example D. The mutant allele $1°2^{\Delta Zn}$ was obtained from Dr. Philip S. Perlman. The mutant allele was constructed by site-directed mutagenesis of pJVM16 and has the following change in the aI2 intron: positions 2028-2036 changed from CCTCATTCA to TAATAATAA ($P_{676}H_{677}S_{678}$ to OchOchOch). The RNP preparation of comparative example D contains a modified, excised aI2 RNA and an aI2-encoded protein that lacks the Zn domain.

Preparation of Substrate DNA

For methods in which it is desired that the nucleotide integrase function as a tool for cleaving double-stranded DNA substrate, it is highly preferred that the DNA substrate have a recognition site that is complementary to the hybridizing region of the group II intron RNA of the nucleotide integrase. When such recognition site is not present in the intended substrate, it is preferably added to the substrate. Examples of substrates to which a recognition site has been added are described below. Of course, where the nucleotide integrase is being used to characterize DNA substrates or to determine if the DNA contains

-18-

certain target sequences, a recognition site is not added to the DNA substrate.

Plasmid pE2E3

5   Plasmid pE2E3 was made by cloning a 71 base pair insert consisting of the yeast mitochondrial COX1 exons 2 and 3 into the SmaI site of pBluescript KS+, obtained from Stratagene, La Jolla, CA to produce a 3032 base pair plasmid containing the target sequence 5'TTTTAGTAGCTGGTCATGCTGTATTAATAATTTTCTTCTTAGTAATGCCTGCT TTAATAGGAGGTTTTGGT), SEQ. ID. NO. 5, in the sense strand. The insert
10 was generated from the mitochondrial DNA of yeast strain 1°2° by PCR using the primers E2-GII-O (5'TTTTAGTAGCTGGTCAGCTGTATT), SEQ. ID. NO. 10, and E3-GII-O (5'ACCAAAACCTCCTATTAAAGCAGGC), SEQ. ID. NO. 11. The insert differs from wild-type 161 sequence at three positions: T to C at position 10 of exon 2, G to A at position 29 of exon 2, and T to A
15 at position 23 of exon 3.

Plasmid pE1E2

Plasmid pE1E2 was made by cloning a 105 base pair insert consisting of a portion of exon 1 and all of exon 2 of the yeast mitochondrial COX1 gene into the SmaI site of pBluescript KS+, obtained
20 from Stratagene to produce a 3066 base pair plasmid. The insert was generated from the mitochondrial DNA of yeast strain m161/m161 C1036 delta 1 delta 2 whose construction is described in Kennell et al., 1993, by using the primers EIJY (5'TAATCATTAGATTAGAATTAGCTGCACCTG), SEQ.ID.NO. 8 and E2JY (5'AGAAAATCATTAATACAGC), SEQ.ID. NO. 9.

25 Plasmid pJVM159

Plasmid pJVM159 was constructed as described in Kennell et al., 1993, Cell 73, 133-146. pJVM159 was obtained by cloning the 4.5 kb fragment of the yeast mitochondrial COX1 gene of strain 1°2° from an MspI site 217 nucleotides upstream of the start codon of exon 1 through
30 the BamHI site in intron aI3 into pBluescript KS (+).

Double-Stranded 142 Base-Pair DNA Substrate from pE2E3

A 142 base pair double-stranded DNA substrate containing exons 2 and 3 of the yeast mitochondrial COX1 was generated from pE2E3 by PCR with the primers KS and SK. To prepare internally-labeled substrate,
35 PCR was carried out in 25 µl of reaction medium containing 1 ng plasmid, 100 ng of each primer, 50 µCi[α-^{32}P]-dTTP (3,000 Ci/mmole) from DuPont NEN, Boston, MA, 30 µM dTTP, 200 µM each of dATP, dCTP and dGTP and 2.5 units Taq DNA polymerase from Life Technologies for 25 cycles (94°C, 30 seconds; 55°C, 30 seconds; 72°C, 30 seconds). After PCR, the
40 internally-labeled product was ethanol precipitated and purified in a 1.5% agarose gel, extracted with phenol-chloroform-isoamyl alcohol having ratio of 25:24:1 and then ethanol precipitated.

Double-Stranded 141 Base-Pair DNA Substrate from pE2E3

A similar 141 base pair internally-labeled double-stranded DNA substrate having a recognition site was synthesized from pE2E3 using primers T7 (5'GTAATACGACTCACTATAGGGC), SEQ. ID. NO. 10, and HG3

5   (5'CAAAAGCTGGGTACCGGGCCCCCCC), SEQ. ID. NO. 11. The PCR reactions were carried out as in the above paragraph. The amplified DNA substrate was digested with *NotI* and *XhoI* from Life Technologies, Inc. to remove 3' end heterogeneity resulting from PCR. The amplified DNA substrate was then ethanol precipitated and purified in a 1.5% agarose gel containing

10   90 mM Tris-borate, pH 8.3, 2 mM EDTA to provide a pE2E3-generated, internally-labeled, 141 base pair, double-stranded, linear DNA substrate.

5' End-labeled DNA Substrates from pE2E3

A 142 base pair double-stranded DNA substrate that was labeled

15   on the 5' end of the sense strand was generated from pE2E3 by PCR with end-labeled primer SK (5'CGCTCTAGAACTAGTGGATC), SEQ.ID.NO.7, and unlabeled primer KS, both of which are complementary to a sequence in the polylinker. 0.2 μg of the primer SK was radiolabeled by using T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and 150 μCi

20   [γ-^{32}P]-ATP (3,000 Ci/mmole; DuPont (NEN). The PCR was carried out as in the above paragraph except in 100 μl of reaction medium with 4 ng plasmid template, 200 ng 5' end-labeled primer and 300 ng unlabeled primer. Following PCR, the labeled substrate DNAs were purified in a 1.5% agarose gel, extracted with phenol-chloroform-isoamyl alcohol

25   (phenol-CIA; 25:24:1) and ethanol precipitated.

A 142 base pair double-stranded DNA substrate that was labeled on the 5' end of the antisense strand was generated from pE2E3 by PCR with 200 ng of the 5' end-labeled primer KS (5'TCGAGGTCG ACGGTATC), SEQ.ID.NO.6, and unlabeled primer SK, both of which are

30   complementary to a sequence in the polylinker. 0.2 μg of primer KS was radiolabeled by using T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and 150 μCi [γ-^{32}P]-ATP (3,000 Ci/mmole; DuPont (NEN). The PCR was carried out as in the above paragraph. Following PCR, the labeled substrate DNAs were purified in a 1.5% agarose gel, extracted

35   with phenol-chloroform-isoamyl alcohol (phenol-CIA; 25:24:1) and ethanol precipitated.

3' End-labeled DNA Substrates from pE2E3

To prepare 3' end-labeled substrates, 1.25 μg of unlabeled PCR product generated from pE2E3 with primers SK and KS was digested with

40   *NotI* or *EcoRI*, and the recessed 3' ends were filled in with [α-^{32}P]-dCTP or [α-^{32}P]-dTTP, respectively, and other unlabeled dNTPs by using Klenow DNA polymerase from Life Technologies, Inc. To ensure that only the desired 3' end was radiolabeled, the opposite end was then digested with *EcoRI* or *BamHI*.

-20-

<u>Double-Stranded DNA Substrates from pE1E2</u>

A 176 base pair internally-labeled double-stranded DNA substrate was synthesized from pE1E2 using primers SK and KS. The PCR reactions were carried out as for the internally labeled 142 basepan double-
5   stranded DNA substrate. The amplified DNA substrate was ethanol precipitated and purified in a 1.5% agarose gel containing 90 mM Tris-borate, pH 8.3, 2 mM EDTA to provide a pE1E2-generated, double-stranded, linear DNA substrate.

<u>5' End-Labeled Double-Stranded DNA Substrate From pE1E2</u>

10 A 176 base pair double-stranded DNA substrate that was labeled on the 5' end of the sense strand was generated from pE1E2 by PCR with 200 ng of the primer SK that was 5' end-labeled using phage T4 polynucleotide kinase and $[\gamma^{-32}P]$-ATP (3,000 Ci/mmole; DuPont NEN). The PCR was carried out as in the above paragraph except in 100 μl of
15 reaction medium with 4 ng plasmid template, 200 ng 5' end-labeled primer and 300 ng unlabeled primer. A 176 base pair double-stranded DNA substrate that was labeled on the 5' end of the antisense strand was generated from pE1E2 by PCR with 200 ng of the primer KS that was 5' end-labeled using phage T4 polynucleotide kinase and $[\gamma^{-32}P]$-ATP
20 (3,000 Ci/mmole; DuPont NEN). The PCR was carried out as above in 100 μl of reaction medium with 4 ng plasmid template, 200 ng 5' end-labeled primer and 300 ng unlabeled primer SK.

<u>Single-Stranded DNA Substrate from pE2E3</u>

A 71 base-pair double-stranded DNA substrate having the
25 recognition site was synthesized from plasmid pE2E3 by PCR with the oligonucleotides E2-GII-O (5'TTTTAGTAGCTGGTCAGCTGTATT), SEQ. ID. NO. 12, and E3-GII-O (5'ACCAAAACCTCCTATTAAAGCAGGC), SEQ. ID. NO. 13. The single stranded DNA substrate has the sequence 5'TTTTAGTAGCTGGTCATGCTGTATTAATAATTTTCTTCTTAGTAATGCCTGCTTTAAT
30 AGGAGGTTTTGGT), SEQ. ID. NO. 5, which is identical to the sense-strand of the double-stranded DNA substrate. The oligonucleotide was 3' end labeled with $[\alpha^{-32}P]$-dTTP using terminal transferase from Life Technologies, Inc., according to the manufacturer's protocol. The labeled oligonucleotide was purified on a G-25 (Sigma) spin column,
35 extracted with phenol-chloroform-isoamyl alcohol (phenol-CIA; 25:24:1), and ethanol precipitated. The specific activity was then adjusted to that of the double stranded DNA substrate by addition of unlabeled oligonucleotide.

<u>Single-Stranded RNA Substrate</u>

40 A 142 nucleotide RNA substrate was synthesized with phage T7 RNA polymerase from pE2E3 linearized with <i>EcoRV</i>. The RNA transcripts were synthesized in 25 μl of reaction medium containing 40 μCi$[\alpha^{-32}P]$-UTP (3,000 Ci/mmol; DuPont NEN), 30 μM UTP, and 200 μM ATP, CTP and GTP. The RNA substrate was precipitated and purified in a denaturing 5%

polyacrylamide gel. The labeling conditions were formulated so that the 3' exon of the RNA substrate has the same specific activity as the 3' exons of the DNA substrates derived from pE1E2.

Double-Stranded DNA from Neurospora crassa

5   A 180 base pair internally-labeled substrate was generated from *Neurospora crassa* strain 74A, which lacks the *coI* intron. The PCR reaction was carried out in 25 µl of reaction medium containing 1 ng mtDNA from strain 74A, 100 ng of the primer 5'GAGTTAAGCGGACCTGGGGTGCAG, SEQ ID. NO. 14, 100 ng of the primer 5'ATTAAGTCTTGGGAATGCCATGTC, SEQ

10   No. 15, 40 µCi [α^{32}P] dTTP (3000 Ci/mmol; DuPont NEN), 200 µM each of dATP, dCTP and dGTP, 2.5 U Taq DNA polymerase for 25 cycles of 94°C, 30 seconds; 55°C, 30 seconds; 72°C, 30 seconds. After PCR, the product was ethanol precipitated and gel-purified. The resulting substrate had the   sense   strand   sequence   of

15   5'GAGTTAAGCGGACCTGGGGTGCAGTACATTGCAGATAATCAATTATACAA

TGCTATAATAACTGCACATGCGATCTTAATGATTTTCTTTATGGTTATGCCAGCATTAATAGGTGGATT

TGGTAATTTCTTGTTACCATTATTAGTAGGGGGTCCTGACATGGCATTCCCAAGACTTAAT,

Seq.Id.No. 16.

Characterization of the RNP particle.

20   0.3 O.D. $_{260}$ of the RNP particle preparations of examples 1 and 5 were applied to a 1% nondenaturing agarose gel in 90 mM Tris borate, pH 8.3, 2 mM EDTA at 4°C and electrophoresed. The particles were transferred to a positively charged synthetic membrane from ICN by capillary transfer with 20 X SSC which contains 3M NaCl, 0.3 M

25   Na-citrate at, pH 7.0. The membrane was rinsed at 4°C in reaction buffer containing 50 mM Tris, pH 8.5, 100 mM KCl, 2 mM MgCl$_2$, 5 mM DTT with 0.05% Nonidet P-40 and 0.2% bovine serum albumin for 15 minutes. The membrane was then incubated for 15 minutes at 37°C in 4 mL reaction buffer with 0.2% BSA 400 µCi [$\alpha32P$] dCTP (3000 Ci/mmol; DuPont NEN),

30   and 0.2 mM each of dATP, dGTP and dTTP. The reaction was chased with 0.2 mM dCTP for 15 minutes at 37°C. Finally, the membrane was washed 3 times for 15 minutes in 10% trichloroacetic acid, 1% sodium pyrophosphate, and exposed for autoradiography.

The lane of the gel containing RNP particles of example 5

35   demonstrated a single radiolabeled band. The appearance of a single band indicated that the aI2 protein of example 5, which has reverse transcriptase activity, migrates as a particle in the nondenaturing gel in which it is associated with an endogenous RNA molecule which acts as a template. The aI2 protein specifically copies an aI2 RNA sequence

40   and does not efficiently copy nonspecific RNA sequences. Kennell et al., 1993, Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. Cell 73, 133-146. Thus, the presence of a single band shows that the aI2 protein is associated with a specific aI2 RNA molecule. The RNP particles of

-22-

example 1 did not produce a similar radiolabeled band. It is believed
that this is due to an inhibition of endogenous reverse transcriptase
activity in wild-type cells.

5   The RNP particles were resolved on a gel, blotted onto a
membrane and probed with a radiolabeled RNA which is complementary to
the aI2 RNA. This experiment showed that aI2 RNA with the size of
excised aI2 RNA ran on the nondenaturing gel as a single band which co-
migrated with the reverse transcription signal. Thus, the aI2 protein
is associated in RNP particles with excised aI2 RNA.

10   A western blot was also prepared and showed that the aI2 protein
that is present in the RNP particles of examples 1,2,4 5, 6,8, and 9 is
a 62 kDa protein having reverse transcriptase function. Thus, the RNP
particle which is isolated from the mitochondria of yeast strain $1°2^{+t}$
comprises an excised aI2 RNA and an associated 62 kDa aI2-encoded

15   protein.

Cleaving Both Strands of a Double-Stranded DNA Substrate and Inserting
an RNA Molecule into the Cleavage Site of One Strand

        Both strands of a double-stranded DNA substrate were cleaved at
37°C and in a medium containing 100 mM KCl, 20 mM $MgCl_2$, 5 mM DTT, at
20   pH 7.5, unless otherwise noted.

Cleaving both strands of the pE2E3-derived DNA Substrate and Inserting
an RNA Molecule

        465 fmoles (1,950,000 cpm) of the uniformly-labeled, 141 base
pair DNA substrate were reacted with 0.375 O.D. units of the RNP
25   particles of Example 1 in a volume of 150 μl. At 0, 1, 3, 5, 10, 20,
and 30 minutes after reaction, 20 μl portions were removed and the
reaction terminated by adding 80 μl of 10 mM EDTA, 0.3 M NaOAc plus 2
μg linear acrylamide carrier, followed by extraction with phenol-
chloroform-isoamyl alcohol (phenol-CIA; 25:24:1) and precipitation with
30   ethanol. One-half of the precipitated products was treated with 0.1 μg
RNase A from Sigma, St. Louis, MO, in 50 μl of 10 mM Tris-HCl, pH 8.0,
1 mM EDTA for 15 minutes at 37°C. The products that had been treated
with RNase A and the products that had not been treated with RNase A
were then analyzed in a denaturing 6% polyacrylamide gel alongside a
35   sequencing ladder generated from pE2E3 using the 5' end-labeled KS
primer.

        Only three bands were detected for the products that were not
treated with RNAse A. These bands correspond to the 69 nucleotide 5'
fragment of the antisense strand, the 72 nucleotide 3' fragment of the
40   antisense strand and the 63 nucleotide 5' fragment of the sense strand
of the DNA substrate. The missing 78 nucleotide product, corresponding
to the 3' fragment of the sense strand, was detected only in those
samples of DNA reaction products that were treated with RNase A prior
to electrophoresis. These results indicate that the DNA substrate had

-23-

been cleaved into four fragments and that an RNA molecule had been
attached to the 3' fragment of the sense strand during reaction with
the RNP particles.

Cleaving both strands of the pE2E3-derived DNA Substrate with RNP
5   Particle Preparation of Examples 1, 2, 4 and 5

0.025 O.D.$_{260}$ units of the RNP particle preparations of examples
1, 2, 4, 5 and comparative examples A-D were reacted with 125 fmoles
(150,000 cpm) of the 142 base pair internally-labeled DNA substrate
generated from pE2E3, for 20 minutes. To verify cleavage, the products
10 were glyoxalated and analyzed in a 1% agarose gel.

A dark band of radiolabel of approximately 2.5 kb RNA was
detected in the lanes containing the products that resulted from
incubation of the DNA substrate with the RNP particles of Examples 1
and 2. A lighter band of RNA was detected in the lane containing the
15 products that resulted from incubation of the DNA substrate with ·
example 4, indicating that these RNP particles cleaved the sense strand
of the substrate and catalyzed insertion of an RNA molecule into the
cleavage site. The RNP particles of example 4 contain a mutated,
excised aI2 intron RNA and an aI2-encoded protein which has a mutation
20 in the reverse transcriptase domain. An even lighter band was detected
in the lane containing products that resulted from incubation of the
DNA substrate with the RNP particles of example 5, which contain a
mutated, excised aI2 RNA and an aI2-encoded protein with a mutation in
the Zn domain. No bands were detected in the lanes which contained the
25 comparative examples. Thus RNP preparations which lack excised aI2
intron RNA or in which the intron-encoded protein lacks the
nonconserved portion of the Zn domain, will neither cleave the DNA
substrate nor attach an RNA.

Analyzing the DNA-RNA Reaction Products

30   625 fmoles (750,000 cpm) of the internally-labeled 142 base pair
substrate DNA generated from pE2E3 were incubated with 0.125 O.D.$_{260}$
units of the RNP particles of Example 1 in 50 µl of medium containing
100 mM KCl, 20 mM MgCl$_2$, 50 mM Tris-HCl, pH 7.5, and 5 mM DTT at 37°C
for 20 minutes. Then 40 µl of 10 mM EDTA, 0.3 M NaOAc and 2 µg linear
35 acrylamide were added to the incubation mixture. The reaction products
were extracted with phenol-CIA and precipitated with ethanol. The
precipitated reaction products were divided into five portions and four
of the portions were further treated with RNase A, alkali, S1 nuclease
or DNase I. RNase A digestion was with 0.1 µg RNase A (Sigma,
40 St. Louis, MO) in 50 µl of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA for
15 minutes at 37°C. S$_1$ nuclease digestion was with 8 units S$_1$, nuclease
in 50 µl of 30 mM NaOAc, pH 4.6, 50 mM NaCl, 1 mM ZnCl$_2$, and 1 µg
single-stranded DNA (salmon sperm; Sigma) for 30 minutes at 25°C.
DNase I digestion was with 7.5 units DNase I (Pharmacia, Piscataway,

-24-

NJ) in 50 μl of 100 mM NaOAc, pH 4.6, 5 mM MgCl₂ for 30 min at 37°C.
Alkali treatment was with 50 μl of 0.1 N NaOH at 50°C for 1 hour,
followed by ethanol precipitation in the presence of 0.3 mM NaOAc at pH
5.2. The reactions products were then glyoxylated and analyzed in a 1%
5   agarose gel containing 10 mM sodium phosphate, pH 7.0 followed by
autoradiography of the dried gel.

        The gel revealed two closely spaced bands of about 2.5-2.7 kb
RNA in the lane containing the products that resulted from incubation
of the DNA substrate with the RNP particles. These bands were absent
10   from the lanes containing substrate alone. These bands were also
absent from the lanes containing products that had been treated with
nucleases or alkali. Thus the 2.5 kb RNA was attached to the DNA
substrate during the reaction. The finding that the RNA-DNA products
withstand denaturation with glyoxal indicates a covalent linkage.

15   Cleavage of 3' and 5' end labeled DNA

        0.025 O.D.₂₆₀ of the RNP particles of Example 1 were incubated
for 20 minutes with 150,000 cpm of each of the 5' and 3' end-labeled
DNA substrates generated from pE2E3, corresponding to 250 fmoles of 5'
end-labeled substrates and 200 fmoles of 3' end-labeled substrates. To
20   verify cleavage, the products were denatured with glyoxal and analyzed
in a 1% agarose gel.

        Two closely spaced dark bands of RNA were detected only with the
substrate labeled at the 3' end of the sense strand. This indicates
that both bands result from the ligation of the aI2 RNA to the 3'
25   fragment of the sense strand. When the gels were exposed for a longer
time, a lighter band was detected with the substrate that had been
labeled at the 5' end of the sense strand. This indicates that a
portion of the aI2 RNA is fully integrated into the cleavage site of
the sense strand during the 20 minute incubation.

30   To confirm that the RNP particle preparation catalyzed the full
integration of the RNA molecule into the cleavage site of the sense
strand of the substrate, the radiolabeled products were eluted from the
gel, subjected to reverse transcriptase-PCR, and sequenced. The PCR
products included the sequence of exon 2 of the yeast mitochondrial
35   *COX1* gene followed by the sequence of the aI2 intron. Thus, the
nucleotide integrase preparation catalyzed attachment of the excised
aI2 RNA molecule to both fragments of the sense strand , i.e. full
integration of the aI2 RNA molecule into the sense-strand cleavage
site.

40   Identifying the RNA in the RNA-DNA Reaction Product

        To confirm that the aI2 intron RNA is attached to the cleaved
strand of the DNA substrate, oligo-directed Rnase H digestions were
performed. For this analysis, oligonucleotides complementary to aI2
RNA were hybridized to an RNA sequence, and the sample was subjected to

RNase H digestion which digests RNA in a RNA-DNA hybrid, but not single-stranded RNA. 0.025 O.D.$_{260}$ of the products were isolated on a 1% native agarose gel and eluted. After precipitation, the products, corresponding to 0.025 O.D.$_{260}$ starting materials, were combined with 40

5   ng each of one or more oligonucleotides in 10 μl of 40 mM Tris, pH 7.3, 100 mM KCl and 2 mM DTT. The products were heated to 90°C for 2 minutes to denature the RNA and cooled on ice. 1 μl 100 mM MgCl$_2$, 0.5 μl RNasin (5 U) and 0.5 μl RNase H (1.4 U) were added, and the mixture was incubated for 30 minutes at 37°C. After phenol-CIA extraction and

10   precipitation, the samples were glyoxalated and resolved on an agarose gel. The four oligonucleotides were found individually and in combination to result in a shift of migration of the products. In contrast, oligonucleotides which are complementary to exon 1, or to aI3 did not shift the migration of the products. Thus, the aI2 RNA is

15   attached to the DNA substrate during the reaction.

        Separate studies on a polyacrylamide gel, using an in vitro transcript of the aI2 RNA confirmed that both the 2.5 kb RNA-DNA product and the 2.7 kb RNA-DNA product represent two different forms of an aI2 RNA lariat attached to the cleaved fragment of the sense

20   strand of the DNA substrate.

Identifying Nucleic Acid Sequences in Substrate DNA

        The nucleotide integrase is useful to identify the presence of particular target sites in a double stranded substrate DNA or to cleave a double stranded substrate DNA which is known to possess the target

25   site.

        0.025 O.D.$_{260}$ units of the RNP particles of examples 1 and 2 were incubated for 20 minutes with 125 fmoles (150,000 cpm) of 3' end-labeled double-stranded DNA substrate derived from pE2E3 which contains the   target   sequence   5'TTTTAGTAGCTGGTCATGCTGTATTAATAATTTTC

30   TTCTTAGTAATGCCTGCTTTAATAGGAGGTTTTGGT), SEQ. ID. NO. 5, in the sense strand substrates. The products were extracted with CIA-phenol, ethanol-precipitated, glyoxylated and analyzed on 1% agarose gels.

        A dark radiolabeled band of 2.5 kb RNA was detected in the products that resulted from the reaction between the RNP particles and

35   the DNA that contained SEQ. ID. No. 5. A radiolabeled band was not detected in products resulting from incubation of RNP particles with a substrate that lacked the target sequence, or with a substrate lacking the sequence upstream of the putative cleavage site, ie. nucleotides 1-35 or with a substrate which contained the aI2-E3 junction of the COX1

40   gene. Only a faint radiolabeled band was detected when the RNP particles were incubated with a modified substrate which contains a 4 base pair substitution in IBS, or with a modified substrate that contained a 10 base pair insert at the putative cleavage site, i.e. between nucleotides 36 and 37, or with a substrate that lacked

-26-

nucleotides 37-71, or with a substrate that contained the E2-aI1 junction from the *COX1* gene.

Thus the hybridizing IBS sequences in the substrate DNA are necessary for cleavage of the sense strand of the DNA substrate by the
5 nucleotide integrase of examples 1 and 2. These results also indicate that the presence of additional sequences downstream of the cleavage site in the DNA substrate will impair cleavage of the substrate by the RNP particles of examples 1 and 2.

Cleaving both strands of the pE2E3-derived DNA Substrate and Inserting
10 an RNA Molecule into the Cleavage Site of the Sense Strand Using the Purified RNP Particles

125 fmoles (150,000 cpm) of the internally-labeled 142 base pair substrate DNA generated from pE2E3 were incubated with 10 μl of each of the fractions obtained from the sucrose gradient in Example 1a. Taking
15 into account the composition of the fractions, the final reaction medium of 20 μl contained 100 mM KCl, 20 mM MgCl₂, 50 mM Tris-HCl, pH 7.5, and 5 mM DTT. Following a 20 minute reaction at 37°C for 20 minutes, 30 μl of water, 5 μl 0.3 M NaOAc and 5 μg tRNA were added to the fractions. The reaction products were extracted with phenol-CIA
20 and precipitated with ethanol, glyoxalated and analyzed on a 1% agarose gel containing 10 mM sodium phosphate, pH 7.0 followed by autoradiography of the dried gel. The radiolabeled bands at 2.5 kb were quantified with a Molecular Dynamics Phosphorimager.

Radiolabeled bands of 2.5 kb were detected when the DNA
25 substrate was reacted with each of the fractions. Thus, the purified RNP particles of example 1a are used to cleave both strands of a double-stranded DNA substrate and to insert the aI2 intron RNA into the cleavage site.

Cleaving Both Strands of a pE2E3 derived double-stranded DNA substrate
30 with the Reconstituted RNP Particle Preparation and with the RNP Preparation of Example 10

The reconstituted RNP particle preparation of Example 12 was reacted with 250 fmoles (300,000 cpm) of the 142 base pair DNA substrates generated from pE2E3 and which were 5' end-labeled on either
35 the sense strand or the antisense strand for 20 minutes at 37°C. To verify cleavage of both strands of the substrate, the reaction products were extracted with phenol-CIA in the presence of 0.3 M NaOAc and 2 μg single-stranded salmon sperm DNA followed by precipitation with ethanol. DNA reactions products were analyzed in a 6%
40 polyacrylamide/8 M urea gel, alongside dideoxy sequencing ladders generated from pE2E3 using 5' end-labeled primers KS or SK.

Radiolabeled bands of DNA corresponding to the 5' fragment of the sense strand and to the 5' fragment of the antisense strand were detected for the products resulting from reaction of reconstituted

particle preparation and the 5' end-labeled DNA substrates. Thus the reconstituted particle preparation cleaves both strands of the DNA substrate derived from pE2E3. Similar results, i.e. cleavage of both strands, were obtained when the 5' end labeled substrates were
5   incubated with the RNP particle preparation of example 10.

Cleaving both strands of DNA Substrate Derived from pE1E2 and Inserting an RNA Molecule into the Cleavage Site of the Sense Strand

      100 fmoles (150,000 cpm) of the internally-labeled pE1E2-derived substrate were incubated with 0.025 O.D.$_{260}$ of the RNP particle
10   preparations of either Example 1 or Example 2 in a medium containing 50 mM KCl, 10 mM MgCl$_2$, 50 mM Tris-HCl at pH 7.5, and 5 mMDTT for 20 minutes. A portion of the DNA incubation products were digested with nucleases. The incubation products were extracted with phenol-CIA and precipitated with ethanol. The DNA incubation products were
15   glyoxylated and applied to a 1% agarose gel containing 10 mM sodium phosphate at pH 7.0. The gel was electrophoresed and dried. The dried gel was autoradiographed.

      Two radiolabeled bands of approximately 2.5 kb RNA were detected on the gel when the DNA substrate was incubated with the RNP particles
20   of Example 1. These bands were not detected for the incubation products that were treated with RNase A or alkali. These results indicate that the excised aI1 RNA molecule of approximately 2.5 kb, was attached to the 176 base pair DNA substrate during incubation with the RNP particles of example 1. Radiolabled bands were not detected when
25   substrate DNA was incubated with the RNP particles of Example 2.

Cleavage of End Labeled DNA Substrates

      To further characterize the DNA-RNA products, 0.025 O.D.$_{260}$ of the RNP particles of example 1 were incubated with 150,000 cpm of each of the 5' and 3' end-labeled DNA substrates generated from pE1E2,
30   corresponding to 250 fmoles of 5' end-labeled substrates and 200 fmoles of 3' end-labeled substrates. Following a 20 minute incubation, the products were denatured with glyoxal and analyzed in a 1% agarose gel.

      Radiolabeled bands of approximately 2.5 kb RNA were detected when the RNP particles were incubated with the 5' and 3' end-labeled
35   DNA substrates which indicated that a complete integration of the RNA molecule into the cleavage site of the sense strand had occurred. Radiolabeled bands were not detected when the RNP particles of example 1 were incubated with substrate that had been radiolabeled on either the 5' end or the 3' end of the antisense strand of the substrate DNA.

40   Analyzing the RNA-DNA Product formed by reaction of the pE1E2 derived DNA substrate with Nucleotide Integrase

      The products that resulted from reacting an internally-labeled pE1E2-derived substrate with the RNP particles of Example 1 were further analyzed on a denaturing 3.5% polyacrylamide gel (39:1

-28-

acrylamide-bisacrylamate), using an excised aI2 RNA molecule as a marker. Three bands were detected on the autoradiograph of the gel. The location of the two, slower-migrating bands suggested that they represented a radiolabeled 3' sense-strand fragment attached to an aI1
5   lariat RNA molecule, while the location of the third, faster-migrating band was consistent with the migration of a linear molecule with the length of aI1. Since reacting the RNP particles of example 1 with a 5'end labeled substrate derived from pE1E2 also results in this band, it indicates that the reaction results in a complete integration of the
10 aI1 RNA into the cleavage site of the sense strand.

       Next, to confirm that the RNP preparation of example 1 catalyzed the full integration of the RNA molecule into the cleavage site of the sense strand of the substrate, the radiolabeled products were eluted from the gel. The products were subjected to reverse transcriptase-PCR
15 in which the eluted RNA was first reverse transcribed using a primer complementary to aI1 RNA. The resulting cDNA amplified by PCR using one primer complementary to the RNA sequence and one primer complementary to the vector. The PCR products were gel purified, cloned into pKS (+) digested with Sma 1 and sequenced according the
20 method of Sanger. The PCR products included the sequence of exon 1 of the yeast mitochondrial COX1 gene followed by the sequence of the aI1 intron. The PCR products also included the sequence of the aI2 intron followed by the sequence of exon 2 of the yeast mitochondrial COX1 gene. Thus, the nucleotide integrase catalyzed attachment of the RNA
25 molecule to both fragments of the sense strand , i.e. full integration of the RNA molecule into the sense-strand cleavage site.

Cleaving Both Strands of a Double-Stranded DNA Substrate and Inserting RNA into the cleavage site of one DNA strand and a cDNA into the Cleavage Site of the Second Strand

30   The nucleotide integrases that cleave both strands of double stranded DNA and insert an RNA into the cleavage site of one strand are also employed in the method for inserting cDNA in the second strand. The method for inserting the cDNA into the second strand is quite similar to the method for cutting the two strands and inserting an RNA
35 except where a cDNA is desired deoxynucleotides must be present in the reaction mixture.

Cleavage of Substrate pJVM159 and Attachment of a cDNA

       0.025 O.D.$_{260}$ units of the RNP particle preparations of examples 1, 2, 4 and 5 and of the materials of comparative Examples A and B were
40 combined with 1 μg of plasmid pJVM159 in 10 μl of reaction medium. The reaction medium contained 0.2 mM each of dATP, dGTP and dTTP, 10 μCi [α-^{32}P]-dCTP (3,000 Ci/mmole; DuPont NEN, Boston, MA), 100 mM KCl, and 5 mM dithiothreitol, 2 mM MgCl$_2$, and 50 mM Tris-HCl, pH 8.5. The reaction was initiated by addition of the RNP preparations, incubated

for 10 minutes at 37°C, and chased with 0.2 mM dCTP for another 10 minutes. After the chase period, the reactions were terminated by extraction with phenol-CIA (phenol-chloroform-isoamyl alcohol; 25:24:1) in the presence of 0.3 M sodium acetate, pH 7.8, and 5 μg E. coli tRNA

5   carrier (Sigma, St. Louis, MO). Products were ethanol precipitated twice and resolved in 1% agarose gels containing 90 mM Tris-borate, pH 8.3, 2 mM EDTA and 0.05% ethidium bromide. The gels were dried and autoradiographed.

     Two additional radiolabeled bands were detected in the lanes
10   containing the products that resulted from incubation of the substrate DNA and the RNP particles of examples 1 and 2, as compared to the control which lacked substrate DNA. This result indicates that the RNP particles of examples 1 and 2 catalyze the formation of a DNA molecule on the cleaved DNA substrate. Such additional bands were not detected
15   when the substrate was incubated with the material of comparative example A, with the RNP particles of Example 4, with the RNP particles of example 5 or the material of comparative example B. Thus, a nucleotide integrase which lacks an excised group II intron RNA or which contains a group II intron-encoded protein that lacks a reverse
20   transcriptase domain does not catalyze the formation of a cDNA molecule on the cleaved strand.

Cleavage of pE2E3 substrate and attachment of a cDNA

     Four μg of plasmid pE2E3 were reacted with 0.100 O.D.$_{260}$ units of the RNP particle preparation of example 1 in 40 μl of reaction medium
25   containing 0.2 mM each of dATP, dGTP and dTTP, 40 μCi [α-^{32}P]-dCTP (3,000 Ci/mmole; DuPont NEN, Boston, MA), 100 mM KCl, 5 mM dithiothreitol, 5 mM MgCl$_2$, and 50 mM Tris-HCl, pH 7.5. The reaction was initiated by addition of the RNP particle preparations, incubated for 10 minutes at 37°C, and chased with 0.2 mM dCTP for another 10
30   minutes. The reaction products were phenol extracted, precipitated and divided into four parts. One part was incubated for 15 minutes at 37°C in 50 μl of low salt buffer containing 0.1 μg RNase A from Sigma, 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA. Another part was incubated for 30 minutes at 37°C in 50 μl of medium containing 2.7 units RNase H from
35   Life Technologies, Inc., 100 mM KCl, 10 mM MgCl$_2$, 20 mM Tris-HCl, pH 7.5, and 0.1 mM DTT. The products were extracted with phenol-CIA (phenol-chloroform-isoamyl alcohol; 25:24:1) in the presence of 0.3 M sodium acetate at pH 7.8, and 5 μg E. coli tRNA carrier (Sigma, St. Louis, MO), ethanol precipitated twice and resolved in 1% agarose
40   gels containing 90 mM Tris-borate, pH 8.3, 2 mM EDTA and 0.05% ethidium bromide. The gels were dried and autoradiographed.

     RNase A digestion of the ^{32}P-labeled plasmid products carried out under low salt conditions to degrade both single stranded and double-stranded RNA, reduced most of the ^{32}P-labeled products to a single

predominant band, which migrated close to linear pE2E3. This most likely resulted from removal of the RNA template. RNase H, which is specific for RNA-DNA duplexes, also increased the mobility of the ^{32}P-labeled products, which indicates that the template RNA is based paired
5   to DNA.

Dideoxy-sequencing of the products showed that the first nucleotide incorporated was the A residue ten nucleotides downstream of the intron RNA insertion site. i.e., cDNA formation began at the antisense strand cleavage site. The sequencing ladders of the longer
10   cDNAs showed that the 3' OH of the 5' fragment of the antisense strand is used as the primer for the formation of the cDNA and that the cDNA synthesis is initiated just downstream of the intron RNA and extends into domain VI of the intron RNA.

Cleaving One Strand of Double-Stranded DNA With the RNP Particles of
15   Examples 6, 7, 8, and 9

0.025 O.D.$_{260}$ units of the RNP particles from examples 1,2,4,5,6,7,8,9, and of the material of comparative examples A-D were incubated with 250 fmoles (300,000 cpm) of the 142 base pair DNA substrates that were generated from pE2E3 and which were 5' end-labeled
20   on either the sense strand or the antisense strand. For comparison, the 5' end labeled DNA substrate was also incubated the RNP particle preparation of Example 1 that had been boiled prior to incubation, or had been treated with 1 μg of RNase A at 37°C for two minutes prior to incubation, or had been treated with 1 μg of protease K at 37°C for 2
25   minutes prior to incubation. Following a 20 minute incubation at 37°C, and the DNA incubation products were extracted with phenol-CIA in the presence of 0.3 M NaOAc and 2 μg single-stranded salmon sperm DNA followed by precipitation with ethanol. DNA incubation products were analyzed in a 6% polyacrylamide/8 M urea gel, alongside dideoxy
30   sequencing ladders generated from pE2E3 using 5' end-labeled primers KS or SK.

A radiolabeled band corresponding to the 5' fragment was detected when RNP particles of examples 1 and 2 were incubated with substrates that had been labeled on the 5' end of either the sense
35   strand or the antisense strand of the DNA substrate, indicating that these particles cleaved both strands of the DNA substrate. The RNP particles of Example 1 cleaved the sense strand precisely at the exon 2-exon 3 junction of the sense strand. The RNP particles of examples 1 and 2 cleaved the antisense strand 10 base pairs downstream from the
40   sense strand cleavage site. RNP particles of Example 1 that had been treated with protease K, or RNase A, or boiled did not cleave either strand.

Radiolabeled bands were also detected when the RNP particles of Example 4 were incubated with DNA substrates that had been 5' end-

labeled on either the sense strand or antisense strand, indicating that this nucleotide integrase cleaved both strands of DNA substrate. The RNP particles of example 4 contain a modified, excised aI2 RNA and an aI2-encoded protein which lacks detectable reverse transcriptase

5 activity. Although the extent of cleavage of RNP particles of Example 4 is somewhat reduced compared to cleavage with the RNP particle preparation of Example 1, the endonuclease activity of the RNA is present even when the reverse transcriptase activity of the aI2-encoded protein is absent.

10   The radiolabeled bands were detected when the RNP particles of Example 5 were incubated with the DNA substrate that had been labeled on the 5' end of either the sense strand or antisense strand. In quantitative assays normalized by either $O.D._{260}$ or soluble aI2 reverse transcriptase activity, the cleavage activities for the antisense and

15 sense strands by the RNP particles of Example 5 were 6% and 25%, . respectively, of activities of the RNP particles of Example 1.

A radiolabeled band corresponding to the 5' fragment was detected when the DNA substrate labeled on the 5' end of the sense strand was incubated with the RNP particles of example 6, but a band

20 corresponding to the 5' fragment of the sense strand was not detected when the RNP particles of example 6 were incubated with a DNA substrate that had been labeled on the 5' end of the antisense strand. The RNP particles of example 6 contain a modified, excised aI2 intron RNA and an aI2-encoded protein that has an alteration in one of the putative

25 endonuclease motifs. Similar results were obtained with the RNP particles of example 7, which contains a modified, excised aI2 intron RNA and an aI2-encoded protein in which the conserved portion of the Zn domain is absent. Likewise, RNP particles of examples 8 and 9, each of which contains a modified, excised aI2 intron RNA and an aI2-encoded

30 protein in which there is a mutation in the Zn^{2+}-like motif, cleaved the sense strand but not the antisense strand of the DNA substrate. For the RNP particles of examples 6, 7, 8, and 9, the level of sense-strand cleavage was proportional to the amount of RNA-DNA products detected in the agarose gels. These findings indicate that the antisense strand

35 endonuclease activity of the aI2-encoded protein is associated with the Zn domain.

Comparative Example B, which lacks excised aI2 RNA, and Comparative Example D, which contains an aI2-encoded protein that lacks the Zn domain, did not cleave either DNA strand.

40   A radiolabeled band corresponding to the 5' fragment was detected when the reconstituted RNP particle preparation of example 12 was incubated with substrates that had been labeled on the 5' end of either the sense strand or the antisense strand of the DNA substrate.

-32-

These results establish that the reconstituted RNP particle preparation cleaves both strands of the DNA substrate.

Thus, both the catalytic RNA molecule of the nucleotide integrase and the intron-encoded protein are required for cleavage of 5 both strands of the double stranded DNA. Certain modifications in the Zn domain and the X domain of intron-encoded protein disrupt the cleavage of the antisense strand of the nucleotide integrase. It is believed that the excised group II intron of the nucleotide integrase cleaves the first strand of the double-stranded DNA substrate and that 10 the group II intron-encoded protein of the nucleotide integrase cleaves the second strand of the double-stranded DNA substrate.

Cleaving Single-stranded DNA Substrates

The RNP particle preparation of Example 1 was incubated with the following DNA substrates: 71 base pair internally-labeled double-15 stranded DNA derived from pE2E3; the 3' end-labeled 71 nucleotide single-stranded DNA derived from pE2E3; and an 142 nucleotide in vitro transcript of aI2 RNA. The incubation was conducted in a medium either containing 100 mM KCl, 20 mM $MgCl_2$, 50 mM Tris-HCl, pH 7.5, and 5 mM DTT for 20 minutes at 37°C or in the same reaction medium supplemented to 20 100 mM $MgCl_2$ for 1 hour at 37°C. The DNA and RNA incubation products were extracted with phenol-CIA in the presence of 0.3 M NaOAc, pH 5.2 with 5 μg E. coli tRNA carrier from Sigma, and then ethanol precipitated and glyoxylated. The glyoxylated RNA and DNA products were then analyzed by electrophoresis in a 1% agarose gel and 25 quantitated with a Molecular Dynamics phosphorimager.

Dark radiolabeled bands of approximately 2.5 kb and 2.7 kb RNA were detected for the products that resulted from incubation of the RNP particles with the double-stranded and single-stranded DNA substrates, in 20 mM Mg^{2+}. In contrast, only a very faint band was detected for the 30 products that resulted from incubation of the RNP particles with the RNA substrate in 20 mM Mg^{2+}. Time course experiments showed that the relative rates of RNA insertion into the substrate by the RNP particle preparation at 20 mM Mg^{2+} were 32:62:1 for the double-stranded DNA, single stranded DNA and RNA, respectively. Thus, the preferred 35 substrate for the nucleotide integrase is a double-stranded or single-stranded DNA substrate.

When the magnesium ion concentration of the medium was increased to 100 mM, only faint radiolabeled bands were detected for the products that resulted from incubation of the RNP particles with the double-40 stranded and single-stranded DNA substrates, which indicates that the insertion of an RNA molecule into the cleavage site of double-stranded DNA or single- stranded DNA works better at a Mg^{2+} concentration below 100 mM.

-33-

Cleaving Substrate with IBS Recognition Site

        The nucleotide integrase is useful to identify the presence of
particular target sites in a single stranded substrate DNA or to cleave
a single stranded substrate DNA which is known to possess the target
5   site.

        0.025 O.D._{260} units of the RNP particles of example 1 were with
the 3' end-labeled single stranded DNA substrate derived from pE2E3
that contained the target sequence, 5'TTTTAGTAGCTGGTCATGCTGTAT
TAATAATTTTCTTCTTAGTAATGCCTGCTTTAATAGGAGGTTTTGGT), SEQ. ID. NO. 5 and
10 with three single stranded DNA substrates that contained modified
target sequences. Following a 20 minute incubation, the products were
extracted with CIA-phenol, ethanol-precipitated, glyoxylated and
analyzed on 1% agarose gels.

        A dark radiolabeled band of 2.5 kb of RNA was detected the
15 products when the nucleotide integrase and double-stranded DNA
substrate derived from pE2E3 which contains the target sequence.
Similarly, a radiolabeled band was detected when the RNP preparations
were incubated with a single-stranded substrate that contained either
IBS3, nucleotides 31 to 36 of SEQ. ID No. 5 or that contained IBS4,
20 nucleotides 25 to 30 of SEQ. ID NO. 5. In contrast, bands were not
detected when the RNP preparations were incubated with a single-
stranded DNA substrate that lacked both the IBS3 and IBS4 sequence.
Thus the nucleotide integrase of example 1 is useful for cleaving
single stranded DNA that contains the sequence TTTTCT or TTAATAA.

-34-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Lambowitz, Alan M
                Zimmerly, Steven
                Guo, Huatao
                Yang, Jian

   (ii) TITLE OF INVENTION: Nucleotide Integrase Preparation

  (iii) NUMBER OF SEQUENCES: 17

   (iv) CORRESPONDENCE ADDRESS:
       (A) ADDRESSEE: Calfee, Halter & Griswold
       (B) STREET: 800 Superior Avenue
       (C) CITY: Cleveland
       (D) STATE: Ohio
       (E) COUNTRY: USA
       (F) ZIP: 44114

    (v) COMPUTER READABLE FORM:
       (A) MEDIUM TYPE: Floppy disk
       (B) COMPUTER: IBM PC compatible
       (C) OPERATING SYSTEM: PC-DOS/MS-DOS
       (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

   (vi) CURRENT APPLICATION DATA:
       (A) APPLICATION NUMBER:
       (B) FILING DATE:
       (C) CLASSIFICATION:

 (viii) ATTORNEY/AGENT INFORMATION:
       (A) NAME: Golrick, Mary E
       (B) REGISTRATION NUMBER: 34,829
       (C) REFERENCE/DOCKET NUMBER: 22727/00127

   (ix) TELECOMMUNICATION INFORMATION:
       (A) TELEPHONE: (216) 622-8200
       (B) TELEFAX: (216) 241-0816

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 2483 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GTGCGCCGTT TCGCTTAATT TATCACTGTA TTGAAGTGTT AATTGATAAA CATATCTCTG      60

TTTATTCAAT TAATGAAAAC TTTACCGTAT CATTTTGGTT CTGATTATTA GTAGTAACAT     120

ACATAGTATT TAGATACGTA AACCATATGG CTTACCCAGT TGGGGCCAAC TCAACGGGGA     180
```

```
CAATAGCATG CCATAAAAGC GCTGGAGTAA AACAGCCAGC GCAAGGTAAG AACTGTCCGA    240

TGGCTAGGTT AACGAATTCC TGTAAAGAAT GTTTAGGGTT CTCATTAACT CCTTCCCACT    300

TGGGGATTGT GATTCATGCT TATGTATTGG AAGAAGAGGT ACACGAGTTA ACCAAAAATG    360

AATCATTAGC TTTAAGTAAA AGTTGACATT TGGAGGGCTG TACGAGTTCA AATGGAAAAT    420

TAAGAAATAC GGGATTGTCC GAAAGGGGAA ACCCTGGGGA TAACGGAGTC TTCATAGTAC    480

CCAAATTTAA TTTAAATAAA GCGAGATACT TTAGTACTTT ATCTAAATTA AATGCAAGGA    540

AGGAAGACAG TTTAGCGTAT TTAACAAAGA TTAATACTAC GGATTTTTCC GAGTTAAATA    600

AATTAATAGA AAATAATCAT AATAAACTTG AAACCATTAA TACTAGAATT TTAAAATTAA    660

TGTCAGATAT TAGAATGTTA TTAATTGCTT ATAATAAAAT TAAAAGTAAG AAAGGTAATA    720

TATCTAAAGG TTCTAATAAT ATTACCTTAG ATGGGATTAA TATTTCATAT TTAAATAAAT    780

TATCTAAAGA TATTAACACT AATATGTTTA AATTTTCTCC GGTTAGAAGA GTTGAAATTC    840

CTAAAACATC TGGAGGATTT AGACCTTTAA GTGTTGGAAA TCCTAGAGAA AAAATTGTAC    900

AAGAAAGTAT GAGAATAATA TTAGAAATTA TCTATAATAA TAGTTTCTCT TATTATTCTC    960

ATGGATTTAG ACCTAACTTA TCTTGTTTAA CAGCTATTAT TCAATGTAAA AATTATATGC   1020

AATACTGTAA TTGATTTATT AAAGTAGATT TAAATAAATG CTTTGATACA ATTCCACATA   1080

ATATGTTAAT TAATGTATTA AATGAGAGAA TCAAAGATAA AGGTTTCATA GACTTATTAT   1140

ATAAATTATT AAGAGCTGGA TATGTTGATA AAAATAATAA TTATCATAAT ACAACTTTAG   1200

GAATTCCTCA AGGTAGTGTT GTCAGTCCTA TTTTATGTAA TATTTTTTTA GATAAATTAG   1260

ATAAATATTT AGAAAATAAA TTTGAGAATG AATTCAATAC TGGAAATATG TCTAATAGAG   1320

GTAGAAATCC AATTTATAAT AGTTTATCAT CTAAAATTTA TAGATGTAAA TTATTATCTG   1380

AAAAATTAAA ATTGATTAGA TTAAGAGACC ATTACCAAAG AAATATGGGA TCTGATAAAA   1440

GTTTTAAAAG AGCTTATTTT GTTAGATATG CTGATGATAT TATCATTGGT GTAATGGGTT   1500

CTCATAATGA TTGTAAAAAT ATTTTAAACG ATATTAATAA CTTCTTAAAA GAAAATTTAG   1560

GTATGTCAAT TAATATAGAT AAATCCGTTA TTAAACATTC TAAAGAAGGA GTTAGTTTTT   1620

TAGGGTATGA TGTAAAAGTT ACACCTTGAG AAAAAAGACC TTATAGAATG ATTAAAAAAG   1680

GTGATAATTT TATTAGGGTT AGACATCATA CTAGTTTAGT TGTTAATGCC CCTATTAGAA   1740

GTATTGTAAT AAAAATTAAAT AAACATGGCT ATTGTTCTCA TGGTATTTTA GGAAAACCCA   1800

GAGGGGTTGG AAGATTAATT CATGAAGAAA TGAAAACCAT TTTAATGCAT TACTTAGCTG   1860

TTGGTAGAGG TATTATAAAC TATTATAGAT TAGCTACCAA TTTTACCACA TTAAGAGGTA   1920

GAATTACATA CATTTTATTT TATTCATGTT GTTTAACATT AGCAAGAAAA TTTAAATTAA   1980

ATACTGTTAA GAAAGTTATT TTAAAATTCG GTAAAGTATT AGTTGATCCT CATTCAAAAG   2040

TTAGTTTTAG TATTGATGAT TTTAAAATTA GACATAAAAT AAATATAACT GATTCTAATT   2100

ATACACCTGA TGAAATTTTA GATAGATATA AATATATGTT ACCTAGATCT TTATCATTAT   2160
```

```
TTAGTGGTAT TTGTCAAATT TGTGGTTCTA AACATGATTT AGAAGTACAT CACGTAAGAA    2220

CATTAAATAA TGCTGCCAAT AAAATTAAAG ATGATTATTT ATTAGGTAGA ATGATTAAGA    2280

TAAATAGAAA ACAAATTACT ATCTGTAAAA CATGTCATTT TAAAGTTCAT CAAGGTAAAT    2340

ATAATGGTCC AGGTTTATAA TAATTATTAT ACTATTAAAT ATGCGTTAAA TGGAGAGCCG    2400

TATGATATGA AAGTATCACG TACGGTTCGG AGAGGGCTCT TTTATATGAA TGTTATTACA    2460

TTCAGATAGG TTTGCTACTC TAC                                           2483
```

(2) INFORMATION FOR SEQ ID NO:2:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 2448 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (iii) HYPOTHETICAL: NO

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GTGCGCCTCT CAGTGCGTAT ATTTCGTTGA TGCGTCTAGC ATTAGTATTA TGAATCATCA     60

ATAGATACTT AAAACATATG ACTAACTCAG TAGGGGCTAA CTTTACGGGG ACAATAGCAT    120

GTCATAAAAC ACCTATGATT AGTGTAGGTG GAGTTAAGTG TTACATGGTT AGGTTAACGA    180

ACTTCTTACA AGTCTTTATC AGGATTACAA TTTCCTCTTA TCATTTGGAT ATAGTAAAAC    240

AAGTTTGATT ATTTTACGTT GAGGTAATCA GATTATGATT CATTGTTTTA GATAGCACAG    300

GCAGTGTGAA AAAGATGAAG GACCTAAATA ACACAAAAGG AAATACGAAA AGTGAGGGAT    360

CAACTGAAAG AGGAAACTCT TGAGTTGACA GAGGTATAGT AGTACCGAAT ACTCAAATAA    420

AAATGAGATT TTTAAATCAA GTTAGATACT ATTCAGTAAA TAATAATTTA AAAATAGGGA    480

AGGATACCAA TATTGAGTTA TCAAAAGATA CAAGTACTTC GGACTTGTTA GAATTTGAGA    540

AATTAGTAAT AGATAATATA AATGAGGAAA ATATAAATAA TAATTTATTA AGTATTATAA    600

AAAACGTAGA TATATTAATA TTAGCATATA ATAGAATTAA GAGTAAACCT GGTAATATAA    660

CTCCAGGTAC AACATTAGAA ACATTAGATG GTATAAATAT AATATATTTA AATAAATTAT    720

CAAATGAATT AGGAACAGGT AAATTCAAAT TTAAACCCAT GAGAATAGTT AATATTCCTA    780

AACCTAAAGG TGGTATAAGA CCTTTAAGTG TAGGTAATCC AAGAGATAAA ATTGTACAAG    840

AAGTTATAAG AATAATTTTA GATACAATTT TTGATAAAAA GATATCAACA CATTCACATG    900

GTTTTAGAAA GAATATAAGT TGTCAAACAG CAATTTGAGA AGTTAGAAAT ATATTTGGTG    960

GAAGTAATTG ATTTATTGAA GTAGACTTAA AAAAATGTTT TGATACAATT TCTCATGATT   1020

TAATTATTAA AGAATTAAAA AGATATATTT CAGATAAAGG TTTTATTGAT TTAGTATATA   1080

AATTATTAAG AGCTGGTTAT ATTGATGAGA AAGGAACTTA TCATAAACCT ATATTAGGTT   1140
```

```
TACCTCAAGG ATCATTAATT AGTCCTATCT TATGTAATAT TGTAATAACA TTGGTAGATA    1200

ATTGATTAGA AGATTATATT AATTTATATA ATAAAGGTAA AGTTAAAAAA CAACATCCTA    1260

CATATAAAAA ATTATCAAGA ATAATTGCAA AAGCTAAAAT ATTTTCGACA AGATTAAAAT    1320

TACATAAAGA AAGAGCTAAA GGCCCACTAT TTATTTATAA TGATCCTAAT TTCAAGAGAA    1380

TAAAATACGT TAGATATGCA GATGATATTT TAATTGGGGT ATTAGGTTCA AAAAATGATT    1440

GTAAATAAT CAAAAGAGAT TTAAACAATT TTTTAAATTC ATTAGGTTTA ACTATAAATG     1500

AAGAAAAAAC TTTAATTACT TGTGCAACTG AACTACCAGC AAGATTTTTA GGTTATAATA    1560

TTTCAATTAC ACCTTTAAAA AGAATACCTA CAGTTACTAA ACTAATTAGA GGTAAACTTA    1620

TTAGAAGTAG AAATACAACT AGACCTATTA TTAATGCACC AATTAGAGAT ATTATCAATA    1680

AATTAGCTAC TAATGGATAT TGTAAGCATA ATAAAAATGG TAGAATAGGA GTGCCTACAA    1740

GAGTAGGTAG ATGACTATAT GAAGAACCTA GAACAATTAT TAATAATTAT AAAGCGTTAG    1800

GTAGAGGTAT CTTAAATTAT TATAAATTAG CTACTAATTA TAAAAGATTA AGAGAAAGAA    1860

TCTATTACGT ATTATATTAT TCATGTGTAT TAACTTTAGC TAGTAAATAT AGATTAAAAA    1920

CAATAAGTAA AACTATTAAA AAATTTGGTT ATAATTTAAA TATTATTGAA AATGATAAAT    1980

TAATTGCCAA TTTTCCAAGA AATACTTTTG ATAATATCAA AAAAATTGAA AATCATGGTA    2040

TATTTATATA TATATCAGAA GCTAAAGTAA CTGATCCTTT TGAATATATC GATTCAATTA    2100

AATATATATT ACCTACAGCT AAAGCTAATT TTAATAAACC TTGTAGTATT TGTAATTCAA    2160

CTATTGATGT AGAAATACAT CATGTTAAAC AATTACATAG AGGTATATTA AAAGCACTTA    2220

AAGATTATAT TCTAGGTAGA ATAATTACCA TAAACAGAAA ACAAATTCCA TTATGTAAAC    2280

AATGTCATAT TAAAACACAT AAAAATAAAT TTAAAAATAT AGGACCTGGT ATATAAAATC    2340

TATTATTAAT GATACTCAAT ATGGAAAGCC GTATGATGGG AAACTATCAC GTACGGTTTG    2400

GGAAAGGCTC TTTAACACGT GGCAACATAG GTTAATTTGC TATTTCAT              2448
```

(2) INFORMATION FOR SEQ ID NO:3:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 777 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: peptide

      (iii) HYPOTHETICAL: NO

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

        Ala Pro Thr Ser Ala Tyr Ile Ser Leu Met Arg Thr Ala Leu Val Leu
        1   5   10   15

        Trp Ile Ile Asn Arg Tyr Leu Lys His Met Thr Asn Ser Val Gly Ala
                    20   25   30

```
Asn Phe Thr Gly Thr Met Ala Cys His Lys Thr Pro Met Ile Ser Val
        35              40              45

Gly Gly Val Lys Cys Tyr Met Val Arg Leu Thr Asn Phe Leu Gln Val
    50              55              60

Phe Ile Arg Ile Thr Ile Ser Ser Tyr His Leu Asp Met Val Lys Gln
65              70              75              80

Val Trp Leu Phe Tyr Val Glu Val Ile Arg Leu Trp Phe Ile Val Leu
            85              90              95

Asp Ser Thr Gly Ser Val Lys Lys Met Lys Asp Thr Asn Asn Thr Lys
            100             105             110

Gly Asn Thr Lys Ser Glu Gly Ser Thr Glu Arg Gly Asn Ser Trp Val
        115             120             125

Asp Arg Gly Met Val Val Pro Asn Thr Gln Met Lys Met Arg Phe Leu
    130             135             140

Asn Gln Val Arg Tyr Tyr Ser Val Asn Asn Asn Leu Lys Met Gly Lys
145             150             155             160

Asp Thr Asn Ile Glu Leu Ser Lys Asp Thr Ser Thr Ser Asp Leu Leu
            165             170             175

Glu Phe Glu Lys Leu Val Met Asp Asn Met Asn Glu Glu Asn Met Asn
            180             185             190

Asn Asn Leu Leu Ser Ile Met Lys Asn Val Asp Met Leu Met Leu Ala
        195             200             205

Tyr Asn Arg Ile Lys Ser Lys Pro Gly Asn Met Thr Pro Gly Thr Thr
    210             215             220

Leu Glu Thr Leu Asp Gly Met Asn Met Met Tyr Leu Asn Lys Leu Ser
225             230             235             240

Asn Glu Leu Gly Thr Gly Lys Phe Lys Phe Lys Pro Met Arg Met Val
            245             250             255

Asn Ile Pro Lys Pro Lys Gly Gly Met Arg Pro Leu Ser Val Gly Asn
        260             265             270

Pro Arg Asp Lys Ile Val Gln Glu Val Met Arg Met Ile Leu Asp Thr
    275             280             285

Ile Phe Asp Lys Lys Met Ser Thr His Ser His Gly Phe Arg Lys Asn
    290             295             300

Met Ser Cys Gln Thr Ala Ile Trp Glu Val Arg Asn Met Phe Gly Gly
305             310             315             320

Ser Asn Trp Phe Ile Glu Val Asp Leu Lys Lys Cys Phe Asp Thr Ile
            325             330             335

Ser His Asp Leu Ile Ile Lys Glu Leu Lys Arg Tyr Ile Ser Asp Lys
        340             345             350

Gly Phe Ile Asp Leu Val Tyr Lys Leu Leu Arg Ala Gly Tyr Ile Asp
        355             360             365

Glu Lys Gly Thr Tyr His Lys Pro Met Leu Gly Leu Pro Gln Gly Ser
    370             375             380
```

```
Leu Ile Ser Pro Ile Leu Cys Asn Ile Val Met Thr Leu Val Asp Asn
385             390             395                 400

Trp Leu Glu Asp Tyr Ile Asn Leu Tyr Asn Lys Gly Lys Val Lys Lys
                405             410              415

Gln His Pro Thr Tyr Lys Lys Leu Ser Arg Met Ile Ala Lys Ala Lys
            420             425             430

Met Phe Ser Thr Arg Leu Lys Leu His Lys Glu Arg Ala Lys Gly Pro
            435             440             445

Thr Phe Ile Tyr Asn Asp Pro Asn Phe Lys Arg Met Lys Tyr Val Arg
        450             455             460

Tyr Ala Asp Asp Ile Leu Ile Gly Val Leu Gly Ser Lys Asn Asp Cys
465             470             475              480

Lys Met Ile Lys Arg Asp Leu Asn Asn Phe Leu Asn Ser Leu Gly Leu
            485             490             495

Thr Met Asn Glu Glu Lys Thr Leu Ile Thr Cys Ala Thr Glu Thr Pro
            500             505             510

Ala Arg Phe Leu Gly Tyr Asn Ile Ser Ile Thr Pro Leu Lys Arg Met
            515             520             525

Pro Thr Val Thr Lys Thr Ile Arg Gly Lys Thr Ile Arg Ser Arg Asn
    530             535             540

Thr Thr Arg Pro Ile Ile Asn Ala Pro Ile Arg Asp Ile Ile Asn Lys
545             550             555              560

Leu Ala Thr Asn Gly Tyr Cys Lys His Asn Lys Asn Gly Arg Met Gly
            565             570             575

Val Pro Thr Arg Val Gly Arg Trp Thr Tyr Glu Glu Pro Arg Thr Ile
            580             585             590

Ile Asn Asn Tyr Lys Ala Leu Gly Arg Gly Ile Leu Asn Tyr Tyr Lys
        595             600             605

Leu Ala Thr Asn Tyr Lys Arg Leu Arg Glu Arg Ile Tyr Tyr Val Leu
    610             615             620

Tyr Tyr Ser Cys Val Leu Thr Leu Ala Ser Lys Tyr Arg Leu Lys Thr
625             630             635              640

Met Ser Lys Thr Ile Lys Lys Phe Gly Tyr Asn Leu Asn Ile Ile Glu
            645             650             655

Asn Asp Lys Leu Ile Ala Asn Phe Pro Arg Asn Thr Phe Asp Asn Ile
            660             665             670

Lys Lys Ile Glu Asn His Gly Met Phe Met Tyr Met Ser Glu Ala Lys
        675             680             685

Val Thr Asp Pro Phe Glu Tyr Ile Asp Ser Ile Lys Tyr Met Leu Pro
    690             695             700

Thr Ala Lys Ala Asn Phe Asn Lys Pro Cys Ser Ile Cys Asn Ser Thr
705             710             715              720

Ile Asp Val Glu Met His His Val Lys Gln Leu His Arg Gly Met Leu
            725             730             735
```

-40-

```
Lys Ala Thr Lys Asp Tyr Ile Thr Gly Arg Met Ile Thr Met Asn Arg
         740                 745                 750

Lys Gln Ile Pro Leu Cys Lys Gln Cys His Ile Lys Thr His Lys Asn
         755                 760                 765

Lys Phe Lys Asn Met Gly Pro Gly Met
         770                 775
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 785 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ala Pro Phe Arg Leu Ile Tyr His Cys Ile Glu Val Leu Ile Asp Lys
1             5                 10                15

His Ile Ser Val Tyr Ser Ile Asn Glu Asn Phe Thr Val Ser Phe Trp
         20                25                30

Phe Trp Leu Leu Val Val Thr Tyr Met Val Phe Arg Tyr Val Asn His
         35                40                45

Met Ala Tyr Pro Val Gly Ala Asn Ser Thr Gly Thr Met Ala Cys His
         50                55                60

Lys Ser Ala Gly Val Lys Gln Pro Ala Gln Gly Lys Asn Cys Pro Met
65                70                75                80

Ala Arg Leu Thr Asn Ser Cys Lys Glu Cys Leu Gly Phe Ser Leu Thr
         85                90                95

Pro Ser His Leu Gly Ile Val Ile His Ala Tyr Val Leu Glu Glu Glu
         100               105               110

Val His Glu Leu Thr Lys Asn Glu Ser Leu Ala Leu Ser Lys Ser Trp
         115               120               125

His Leu Glu Gly Cys Thr Ser Ser Asn Gly Lys Leu Arg Asn Thr Gly
         130               135               140

Leu Ser Glu Arg Gly Asn Pro Gly Asp Asn Gly Val Phe Met Val Pro
145               150               155               160

Lys Phe Asn Leu Asn Lys Ala Arg Tyr Phe Ser Thr Leu Ser Lys Leu
         165               170               175

Asn Ala Arg Lys Glu Asp Ser Leu Ala Tyr Leu Thr Lys Ile Asn Thr
         180               185               190

Thr Asp Phe Ser Glu Leu Asn Lys Leu Met Glu Asn Asn His Asn Lys
         195               200               205

Thr Glu Thr Ile Asn Thr Arg Ile Leu Lys Leu Met Ser Asp Ile Arg
```

```
        210                    215                      220
Met Leu Leu Ile Ala Tyr Asn Lys Ile Lys Ser Lys Lys Gly Asn Met
225                 230                 235                 240

Ser Lys Gly Ser Asn Asn Ile Thr Leu Asp Gly Ile Asn Ile Ser Tyr
                245                 250                 255

Leu Asn Lys Leu Ser Lys Asp Ile Asn Thr Asn Met Phe Lys Phe Ser
            260                 265                 270

Pro Val Arg Arg Val Glu Ile Pro Lys Thr Ser Gly Gly Phe Arg Pro
            275                 280                 285

Leu Ser Val Gly Asn Pro Arg Glu Lys Ile Val Gln Glu Ser Met Arg
        290                 295                 300

Met Met Leu Glu Ile Ile Tyr Asn Asn Ser Phe Ser Tyr Tyr Ser His
305                 310                 315                 320

Gly Phe Arg Pro Asn Leu Ser Cys Leu Thr Ala Ile Ile Gln Cys Lys
                325                 330                 335

Asn Tyr Met Gln Tyr Cys Asn Trp Phe Ile Lys Val Asp Leu Asn Lys
            340                 345                 350

Cys Phe Asp Thr Ile Pro His Asn Met Leu Ile Asn Val Leu Asn Glu
            355                 360                 365

Arg Ile Lys Asp Lys Gly Phe Met Asp Leu Leu Tyr Lys Leu Leu Arg
    370                 375                 380

Ala Gly Tyr Val Asp Lys Asn Asn Asn Tyr His Asn Thr Thr Leu Gly
385                 390                 395                 400

Ile Pro Gln Gly Ser Val Val Ser Pro Ile Leu Cys Asn Ile Phe Leu
                405                 410                 415

Asp Lys Leu Asp Lys Tyr Leu Glu Asn Lys Phe Glu Asn Glu Phe Asn
            420                 425                 430

Thr Gly Asn Met Ser Asn Arg Gly Arg Asn Pro Ile Tyr Asn Ser Leu
            435                 440                 445

Ser Ser Lys Ile Tyr Arg Cys Lys Leu Leu Ser Glu Lys Leu Lys Leu
    450                 455                 460

Ile Arg Leu Arg Asp His Tyr Gln Arg Asn Met Gly Ser Asp Lys Ser
465                 470                 475                 480

Phe Lys Arg Ala Tyr Phe Val Arg Tyr Ala Asp Asp Ile Ile Ile Gly
                485                 490                 495

Val Met Gly Ser His Asn Asp Cys Lys Asn Ile Leu Asn Asp Ile Asn
            500                 505                 510

Asn Phe Leu Lys Glu Asn Leu Ser Ile Val Met Lys Leu Asn Lys His
            515                 520                 525

Gly Tyr Cys Ser His Gly Ile Leu Gly Lys Pro Gly Met Ser Ile Asn
    530                 535                 540

Met Asp Lys Ser Val Ile Lys His Ser Lys Glu Gly Val Ser Phe Leu
545                 550                 555                 560

Gly Tyr Asp Val Lys Val Thr Pro Trp Glu Lys Arg Pro Tyr Arg Met
```

-42-

	565			570			575	

Ile Lys Lys Gly Asp Asn Phe Ile Arg Val Arg His His Thr Ser Leu
          580   585   590

Val Val Asn Ala Pro Ile Arg Arg Gly Val Gly Arg Leu Ile His Glu
       595   600   605

Glu Met Lys Thr Ile Leu Met His Tyr Leu Ala Val Gly Arg Gly Ile
    610   615   620

Met Asn Tyr Tyr Arg Leu Ala Thr Asn Phe Thr Thr Leu Arg Gly Arg
625   630   635   640

Ile Thr Tyr Ile Leu Phe Tyr Ser Cys Cys Leu Thr Leu Ala Arg Lys
          645   650   655

Phe Lys Leu Asn Thr Val Lys Lys Val Ile Leu Lys Phe Gly Lys Val
          660   665   670

Leu Val Asp Pro His Ser Lys Val Ser Phe Ser Ile Asp Asp Phe Lys
          675   680   685

Ile Arg His Lys Met Asn Met Thr Asp Ser Asn Tyr Thr Pro Asp Glu
     690   695   700

Ile Leu Asp Arg Tyr Lys Tyr Met Leu Pro Arg Ser Leu Ser Leu Phe
705   710   715   720

Ser Gly Ile Cys Gln Ile Cys Gly Ser Lys His Asp Leu Glu Val His
          725   730   735

His Val Arg Thr Leu Asn Asn Ala Ala Asn Lys Ile Lys Asp Asp Tyr
          740   745   750

Leu Leu Gly Arg Met Ile Lys Met Asn Arg Lys Gln Ile Thr Ile Cys
          755   760   765

Lys Thr Cys His Phe Lys Val His Gln Gly Lys Tyr Asn Gly Pro Gly
    770   775   780

Leu
785

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 71 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTTTAGTAGC TGGTCATGCT GTATTAATAA TTTTCTTCTT AGTAATGCCT GCTTTAATAG   60

GAGGTTTTGG T   71

(2) INFORMATION FOR SEQ ID NO:6:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 17 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCGAGGTCGA CGGTATC   17

(2) INFORMATION FOR SEQ ID NO:7:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 20 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCTCTAGAA CTAGTGGATC   20

(2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TAATCATTAG ATTAGAATTA GCTGCACCTG   30

(2) INFORMATION FOR SEQ ID NO:9:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

-44-

```
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGAAAATCAT TAATACAGC                                                    19

(2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (iii) HYPOTHETICAL: NO




        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTAATACGAC TCACTATAGG GC                                                22

(2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 25 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (iii) HYPOTHETICAL: NO




        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAAAAGCTGG GTACCGGGCC CCCCC                                             25

(2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: cDNA

        (iii) HYPOTHETICAL: NO




        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTTTAGTAGC TGGTCAGCTG TATT                                             24

(2) INFORMATION FOR SEQ ID NO:13:
```

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 25 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACCAAAACCT CCTATTAAAG CAGGC   25

(2) INFORMATION FOR SEQ ID NO:14:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 24 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GAGTTAAGCG GACCTGGGGT GCAG   24

(2) INFORMATION FOR SEQ ID NO:15:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 24 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATTAAGTCTT GGGAATGCCA TGTC   24

(2) INFORMATION FOR SEQ ID NO:16:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 180 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: double
         (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAGTTAAGCG GACCTGGGGT GCAGTACATT GCAGATAATC AATTATACAA TGCTATAATA   60

ACTGCACATG CGATCTTAAT GATTTTCTTT ATGGTTATGC CAGCATTAAT AGGTGGATTT   120

GGTAATTTCT TGTTACCATT ATTAGTAGGG GGTCCTGACA TGGCATTCCC AAGACTTAAT   180

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 105 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TAATCATTAG ATTAGAATTA GCTGCACCTG GTTCACAATA TTTACATGGT AATTCACAGT   60

TATTTAATGT TTTAGTAGTT GGTCATGCTG TATTAATGAT TTTCT   105

-47-

What is Claimed is:

        1. A method for cleaving a double-stranded or single-stranded
DNA substrate having a recognition site, comprising the following
steps:
        (a)   providing a nucleotide integrase comprising:
               (i)   an excised group II intron RNA having a hybridizing
                      region with a sequence complementary to the
                      recognition site on at least one strand of the DNA
                      substrate; and
               (ii)   a protein bound to said RNA and having an amino
                      acid sequence encoded by a group II intron;
        (b)   reacting said nucleotide integrase with the DNA substrate
in the presence of a sufficient concentration of divalent cation so
that at least one strand of the DNA substrate is cleaved and a nucleic
acid molecule is attached to one or both of the cleaved strands.

        2.   The method of claim 1 wherein:
        the DNA is double-stranded;
        said nucleotide integrase is an RNP particle;
        the group II RNA further comprises six domains;
        only one strand of the DNA substrate is cleaved;
        and a nucleic acid molecule is an RNA molecule which is attached
to the cleaved strand.

        3.   The method of claim 1 wherein the DNA is double-stranded;
said nucleotide integrase is an RNP particle; the protein comprises: an
X domain;
        and a Zn domain; each strand of the DNA substrate is cleaved;
and
        a nucleic acid molecule is attached to at least one of the
 cleaved strands.

        4.   The method of claim 3 wherein the group II RNA further
comprises six domains.

        5.   The method of claim 3 wherein a nucleic acid molecule
attached to the cleaved DNA strand is a group II intron RNA.

        6.   The method of claim 3 further comprising the step of:
        reacting deoxynucleotides with the DNA substrate in the
 presence of said nucleotide integrase;

wherein said protein further comprises a reverse transcriptase domain, a nucleic acid molecule is attached to each of the cleaved strands.

7.   The method of claim 6 wherein a group II intron RNA is attached to one strand and a cDNA is attached to the other strand of the DNA substrate.

8.   The method of claim 1 wherein:
the nucleotide integrase is an RNP particle;
the group II intron RNA is a yeast mitochondrial group II intron RNA and the protein amino acid sequence is encoded by a yeast mitochondrial group II intron.

9.   The method of claim 1 wherein: the nucleotide integrase is an RNP particle; the RNA is a fungal mitochondrial group II intron RNA; and the amino acid sequence is encoded by a fungal mitochondrial group II intron.

10.   The method of claim 1 wherein the divalent cation is a magnesium ion at a concentration from about 2 mM to about 20 mM.

11.   The method of claim 1 wherein the divalent cation is a magnesium ion at a concentration less than about 100 mM.

12.   The method of claim 1 wherein the divalent cation is a magnesium ion at a concentration greater than about 1 mM.

13.   The method of claim 1 wherein: said nucleotide integrase is an RNP particle; the RNA is an excised RNA transcript encoded by the group II intron aI2 of the yeast mitochondrial COX1 gene; and the protein is encoded by the group II intron aI2 of the yeast mitochondrial COX1 gene.

14.   The method of claim 1 wherein: said nucleotide integrase is an RNP particle; the RNA is an excised RNA transcript encoded by the group II intron aI1 of the yeast mitochondrial COX1 gene; and the protein is encoded by the group II intron of the aI1 of the yeast mitochondrial COX1 gene.

15.   The method of claim 1 wherein: said nucleotide integrase is an RNP particle; the RNA is an excised RNA transcript encoded by the fungal mitochondrial group II intron coI; and the protein is encoded by the fungal mitochondrial group II intron coI.

16. The method of claim 13 wherein the aI2 intron has the nucleotide sequence of SEQ. ID. NO. 1; said RNA transcript is about 2.5 kb; and said protein has a molecular weight of about 62 kDa.

17. The method of claim 14 wherein the aI1 intron has the nucleotide sequence of SEQ. ID. NO. 3.

18. The method of claim 1 wherein the DNA is double-stranded and the nucleotide integrase is provided in a reconstituted RNP particle preparation which comprises:

    (i)   an exogenous RNA which comprises the excised group II intron RNA having a hybridizing region having a sequence complementary to the recognition site on at least one strand of the DNA substrate; and

    (ii)   an RNA-protein complex comprising:

           the protein having an amino acid sequence encoded by group II intron further comprising, an X domain and a Zn Domain; and RNA which is free of excised group II intron RNA that encodes said protein; and

    (iii) a reconstituted RNP particle comprising said exogenous RNA bound to said protein.

19. The method of claim 18 wherein said exogenous RNA comprises an excised RNA encoded by the group II intron aI2 of the yeast mitochondrial COX1 gene; and said protein is encoded by the group II intron aI2 of the yeast mitochondrial COX1 gene.

20. The method of claim 1 wherein the DNA is single-stranded and the excised group II intron RNA has a hybridizing region complementary to at least one recognition site on the single-stranded DNA substrate.

21. A reconstituted RNP particle preparation, comprising:

(a)   an exogenous RNA which comprises an excised group II intron RNA; and

(b)   an RNA-protein complex comprising:

    (i)   a protein having an amino acid sequence encoded by a group II intron, an X domain and a Zn domain, and

    (ii)   RNA that is free of excised group II intron RNA having a sequence that encodes said protein; and

(c)   a reconstituted RNP particle that cleaves DNA at a specific site, said RNP particle comprising said exogenous RNA bound to said protein.

-50-

22.   The reconstituted RNP particle preparation of claim 21 wherein:

the exogenous RNA comprises a yeast mitochondrial group II intron RNA and the protein comprises an amino acid sequence encoded by a yeast mitochondrial group II intron.

23.   The reconstituted RNP particle preparation of claim 21 wherein:

the exogenous RNA comprises an excised RNA transcript that is encoded by the group II intron aI2 of the yeast mitochondrial *COX1* gene and said protein comprises an amino acid sequence encoded by the group II intron aI2 of the yeast mitochondrial *COX1* gene.

24.   A purified reconstituted RNP particle comprising:

(a) an exogenous RNA which comprises an excised group II intron RNA; and

(b) a protein having an amino acid sequence encoded by a group II intron bound to said RNA;

wherein said RNP particle is substantially free of ribosomes.

25.   A method for making the reconstituted RNP particle preparation of claim 21 comprising the following steps:

(a)   providing an exogenous RNA molecule which comprises an excised group II intron RNA;

(b) providing an RNA-protein complex comprising:

(i)   a protein having an amino acid sequence encoded by a group II intron, an X domain and a Zn domain, and

(ii)   RNA that is free of excised group II intron RNA having a sequence that encodes said protein; and

(c)   combining said exogenous RNA with said RNA-protein complex to form a reconstituted RNP particle.

26.   The method of claim 25 wherein the exogenous RNA comprises a yeast mitochondrial group II intron₁RNA and the protein comprises an amino acid sequence encoded by a yeast mitochondrial group II intron.

27.   The method of claim 25 wherein the exogenous RNA comprises an excised RNA transcript that is encoded by the group II intron aI2 of the yeast mitochondrial *COX1* gene and said protein comprises an amino acid sequence encoded by the group II intron aI2 of the yeast mitochondrial *COX1* gene.

28.   A purified RNP particle, comprising:

-51-

(i)   an excised group II intron RNA having a hybridizing region having a sequence which is complementary to a recognition site on one strand of a DNA substrate;

(ii)   a protein bound to said group II intron RNA, said protein comprising an amino acid sequence which is encoded by a group II intron, an X domain and a Zn domain; and

wherein said RNP particle is substantially free of ribosomal RNA.

29.   The purified RNP particle of claim 28 wherein said group II intron RNA has six domains and said protein further comprises a reverse transcriptase domain.

30.   The purified RNP particle of claim 29 wherein said RNA comprises a yeast mitochondrial group II intron RNA and said protein comprises an amino acid sequence encoded by a yeast mitochondrial group II intron.

31.   The method of claim 1 wherein the nucleotide integrase is reacted with the DNA substrate in the presence of from about 50 mM to about 300 mM KCl and from about 2 mM to about 20 mM Mg^{2+}.
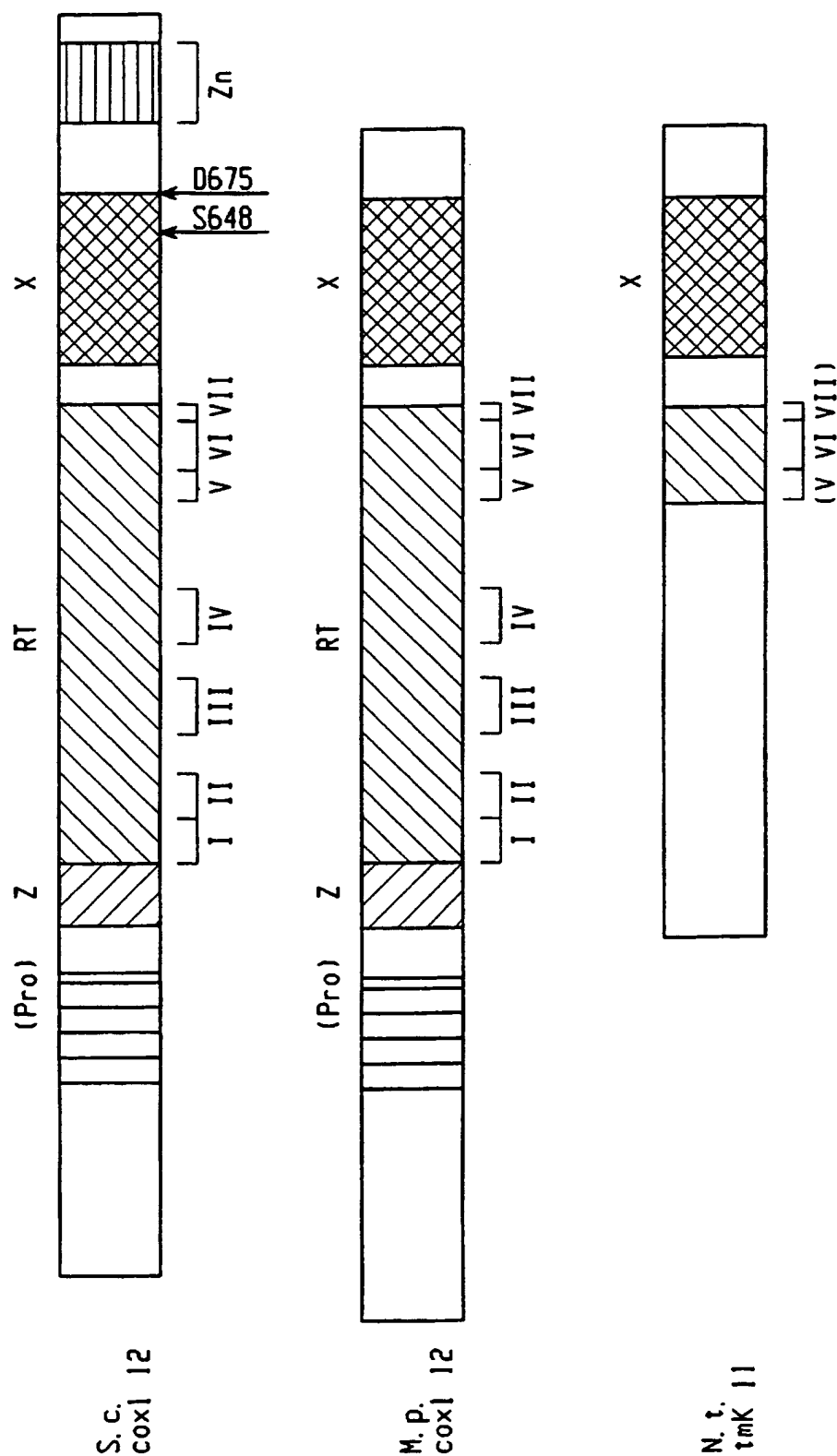
Fig. 1
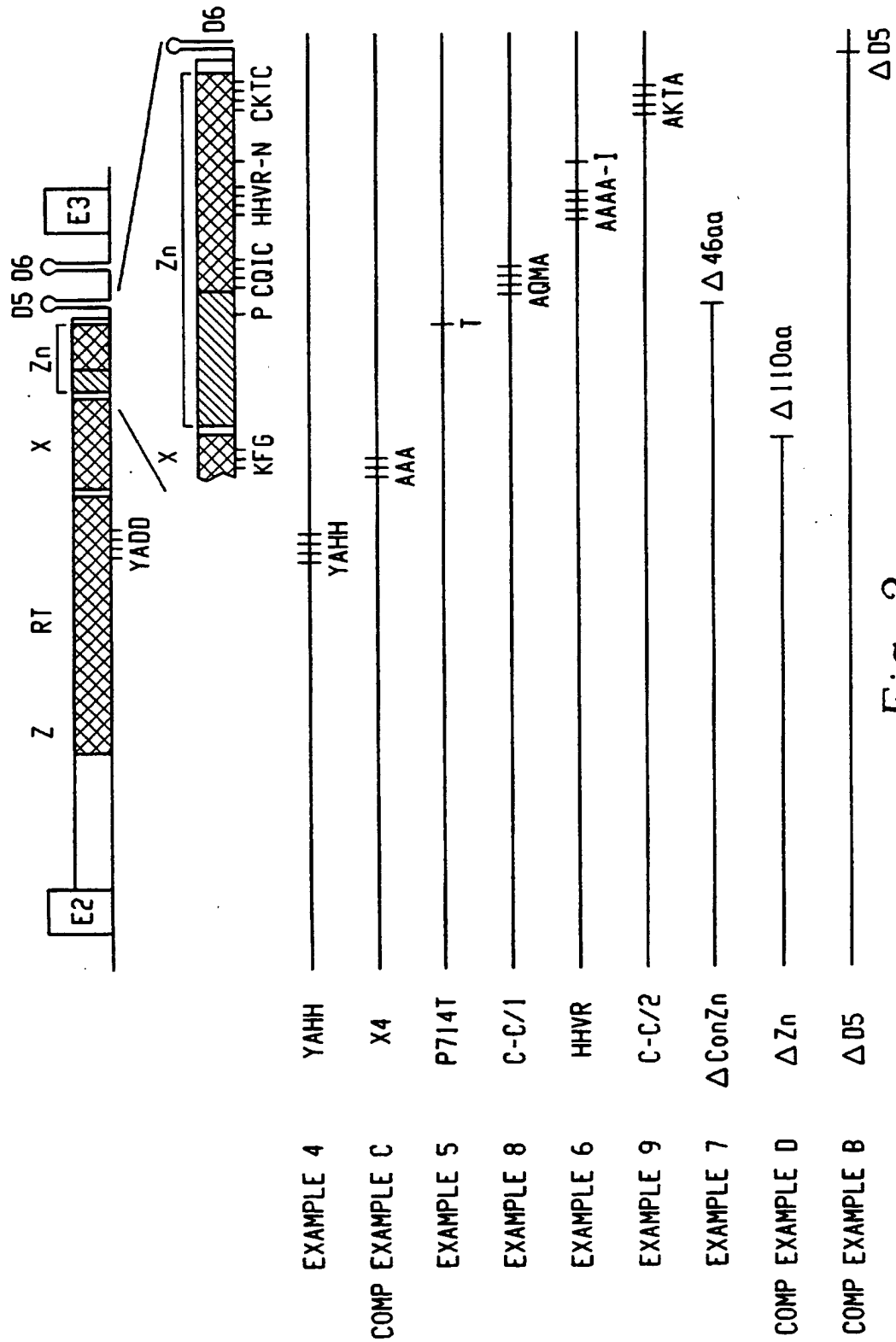
Fig. 2
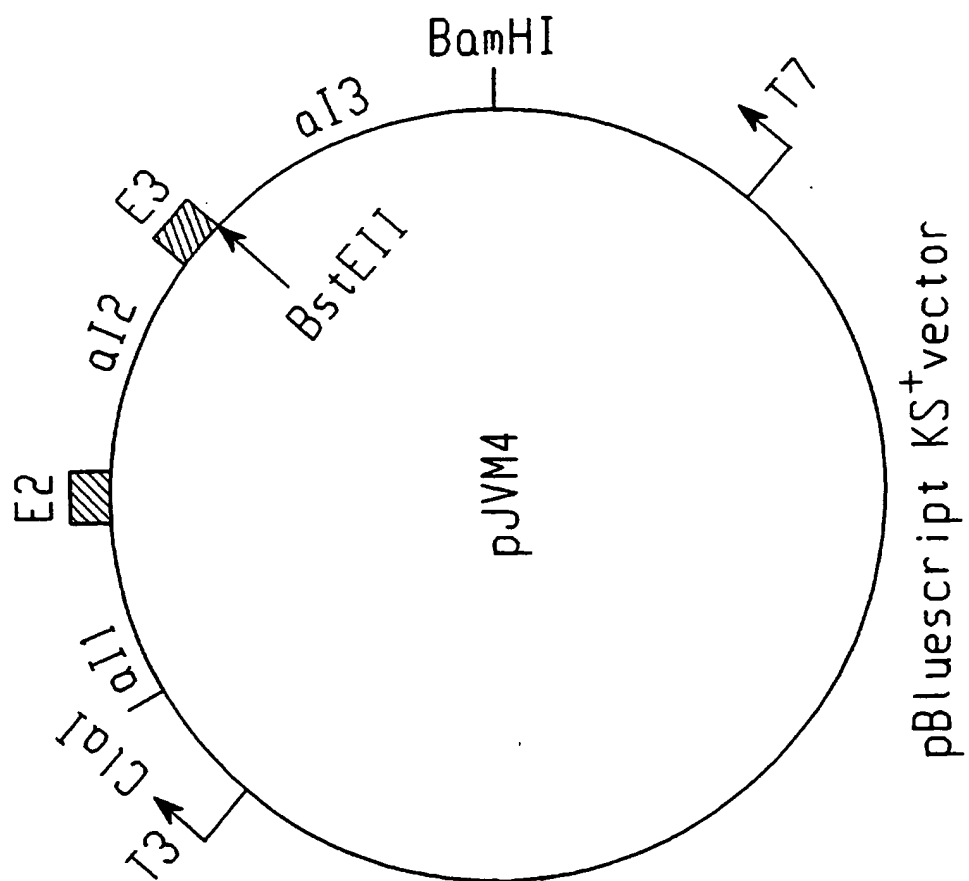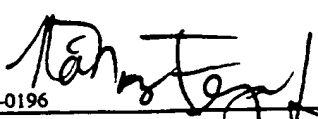
Fig. 3

INTERNATIONAL SEARCH REPORT

A.   CLASSIFICATION OF SUBJECT MATTER

IPC(6)   :C12Q 1/68
US CL   :435/6; 514/44

According to International Patent Classification (IPC) or to both national classification and IPC

B.   FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   435/6, 18; 514/44; 536/24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

BIOSIS, EMBASE, LIFE SCIENCES, MEDLINE, EPO, CA, DERWENT, APS

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,180,818 A (CECH ET AL.) 19 January 1993, columns 1-6, claims 1-8.	1-31
Y	MORL et al. Integration of group II intron bI1 into a foreign RNA by reversal of the self-splicing reaction. Cell. 23 February 1990, Vol. 60, pages 629-636, especially pages 631-634.	1-31
Y	MOHR et al. Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. Nucleic Acids Research. 1993, Vol. 21, No. 22, pages 4991-4997, see entire document.	1-31

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 OCTOBER 1996	06 NOV 1996

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	PAUL B. TRAN, PH.D.
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KENNELL et al. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. Cell. 09 April 1993, Vol. 73, pages 133-146, see entire document.	1-31